# 6.

# CITED2-mediated human hematopoietic stem cell maintenance is critical for acute myeloid leukemia

**PM Korthuis, G Berger, B Bakker, M Geugien, J Jaques, G de Haan, JJ Schuringa, E Vellenga and H Schepers**

*(Leukemia, 2015)*

**Abstract**

As the transcriptional coactivator CITED2 (CBP/p300-interacting-transactivator-with-an ED-rich-tail 2) can be overexpressed in acute myeloid leukemia (AML) cells, we analyzed the consequences of high CITED2 expression in normal and AML cells. CITED2 overexpression in normal $CD34^+$ cells resulted in enhanced hematopoietic stem and progenitor cell (HSPC) output *in vitro*, as well as in better hematopoietic stem cell (HSC) engraftability in NSG mice. This was because of an enhanced quiescence and maintenance of $CD34^+CD38^-$ HSCs, due in part to an increased expression of the cyclin-dependent kinase inhibitor CDKN1A. We demonstrated that PU.1 is a critical regulator of CITED2, as PU.1 repressed CITED2 expression in a DNA methyltransferase 3A/B (DNMT3A/B)-dependent manner in normal $CD34^+$ cells. $CD34^+$ cells from a subset of AML patients displayed higher expression levels of CITED2 as compared with normal $CD34^+$ HSPCs, and knockdown of CITED2 in AML CD34+ cells led to a loss of long-term expansion, both *in vitro* and *in vivo*. The higher CITED2 expression resulted from reduced PU.1 activity and/or dysfunction of mutated DNMT3A/B. Collectively, our data demonstrate that increased CITED2 expression results in better HSC maintenance. In concert with low PU.1 levels, this could result in a perturbed myeloid differentiation program that contributes to leukemia maintenance.

The transcriptional coactivator CBP/p300-interacting-transactivator-with-an ED-rich-tail 2 (CITED2) was originally discovered to bind the CH1 region of CBP/p300.[1] More recently, we showed that CITED2 is essential for adult HSC maintenance.[2] CITED2 deletion in adult HSCs led to rapid animal lethality, resulting from hematopoietic stem and progenitor cell (HSPC)-specific apoptosis. In contrast, lineage-specific deletion of CITED2 did not result in such a dramatic phenotype, suggesting HSC-specific actions of CITED2.[2] These actions might be mediated through control of the INK4a/ARF locus by the PRC1 genes BMI1 and MEL18.[3] In addition, TFAP2A and MYC have been shown to control cellular proliferation of respectively breast and lung cancer cells using CITED2 as a cofactor.[4,5] The fact that CITED2 expression is essential for HSC maintenance and functions as a cofactor in oncogeneinduced transformation suggests that CITED2 could also play a role in leukemia initiation or leukemic stem cell maintenance. Gene expression data from Andersson et al[6] suggest that acute myeloid leukemic (AML) cells can have an enhanced CITED2 expression, while in chronic myeloid leukemia (CML), CITED2 expression is even further increased upon transition into blast crisis.[7,8] How CITED2 expression may be increased during leukemogenesis is unknown, but important transcription factors, like HIF1α,[9] FOXO3A[10] and STAT5,[5] have been shown to regulate its expression. Closer examination of the CITED2 promoter revealed that, besides HIF1α, FOXO3A and STAT5, ETS transcription factor binding sites are also present.

The presence of ETS-binding sites is of interest, since a member of the ETS family of transcription factors, PU.1, is frequently inactivated in AML. Although point mutations of PU.1 are rarely observed,[11,12] many leukemic oncogenes are known to reduce or inactivate PU.1 expression.[11,13-17] Such decreased expression of PU.1 has been shown to induce AML in mice.[18-20] Furthermore, PU.1 haploinsufficiency cooperates with SOX4 to induce AML in mice.[21] Based on these findings we

questioned whether PU.1 modulates CITED2 expression and function.

Here we report that a subset of CD34$^+$ AML cells express increased levels of CITED2. As in normal CD34$^+$ cells, this likely affects quiescence as well as apoptosis of these cells. We furthermore demonstrate that reduced expression or activity of PU.1 contributes to this high expression of CITED2.

## Methods

**Culture conditions.** Long-term cultures on stroma, CFC, LTC-IC and single cell assays were performed as described previously and in the *supplementary methods.*[22,23]

**Flow cytometry analysis.** All fluorescence activated cell sorter (FACS) analyses were performed on a LSRII (*Becton Dickinson (BD), Alphen a/d Rijn, The Netherlands*) or MACSQuant (*Myltenyi Biotec*) and data was analyzed using FlowJo (*Treestar, USA*). Cells were sorted on a MoFLo XDP or Astrios (*DakoCytomation, Carpinteria, CA, USA*). Antibodies were obtained from BD Bioscience (*Breda*), Biolegend (*Uithoorn*), eBioscience (*Vienna*), Milteny (*Leiden*), Nuclilab (*Huissen*) or DAKO (*Enschede*) and are described in the *supplemental methods.*

**In vivo transplantations into NSG mice.** Eight to ten week-old female NSG mice (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ) were purchased from Charles River Laboratory and bred in house. Mouse experiments were performed in accordance with national and institutional guidelines and all experiments were approved by the Institutional Animal Care and Use Committee of the University of Groningen (IACUC-RuG). Experiments are described in the *supplemental methods.*

**(Quantitative) PCR and gene expression profiling.** RNA isolation is described in the *supplementary methods.* Target gene expression was investigated by means of Q-PCR. Typical examples are shown and presented as mean with standard error, sequences can be found in *supplementary table 3.* A detailed description of the gene expression analysis workflow is described in the *supplementary methods.*

**Reporter assays.** For cloning and transfection of constructs see *supplemental methods.*[24,25] Cells were harvested and lysed using Promega Luciferase Cell Culture Lysis Reagent according the manufacturers recommendations. Luciferase signals were measured using a Synergy H4 Hybrid Reader (*BioTek, Germany*). Alternatively, the mean fluorescent intensity of GFP was measured by FACS.

**Chromatin immunoprecipitations.** 2.5 x 10$^6$ CD34$^+$ cells or HL-60 cells were used for chromatin immunoprecipitation studies as described before.[26]

## Results

### CITED2 enhances human HSPC cultures and engraftment

Gene expression profiling suggested that CITED2 expression might be enhanced in AML cells. To study whether high CITED2 expression affects hematopoietic stem and progenitor cell activity, CB CD34$^+$ cells were lentivirally transduced with CITED2 (*Supplementary figures 1A and B*) and hematopoietic development was studied. Improved long-term expansion of hematopoietic cells on MS5 co-cultures was observed upon overexpression of CITED2 (Figure 1A, $p<0.05$), with a marked increase in the size (Figure 1B) and number (*data not shown*) of week 5 cobblestone area forming cells (CAFCs). To investigate the effect of ectopic CITED2 expression on progenitor cells, transduced CD34$^+$ cells were plated in methylcellulose either directly, or after each subsequent week of

culture on MS5 stromal cells. CITED2 overexpression increased the number of colony-forming-cells (CFCs) ~4-fold directly after transduction (Figure 1C). CITED2 increased CFC numbers from both the $CD34^+CD38^+$ and $CD34^+CD38^-$ compartments, but subsequent replating capacity was confined to the $CD34^+CD38^-$ compartment (*data not shown*). This increased progenitor persisted for at least 5 weeks of culture (Figure 1D). These CITED2 expressing colonies were ~5x larger than control colonies (Figure 1E, note magnification), in line with the observed increase in suspension cell numbers. These colonies depicted a clear increase in colony-forming-unitgranulocyte (CFU-G) upon overexpression of CITED2, which was paralleled by a decrease in burst-forming-unit-erythrocytes (BFU-E) (*Supplementary figure 1C*). Similar, in MS5-cocultures an increase in the percentage of $CD15^+$ granulocytic cells was observed (*Supplementary figure 1D*, upper panel), in contrast to a decrease in $CD14^+$ monocytes/macrophages and $GPA^+$ erythroid cells (*Supplementary figure 1D*, upper and lower panels).

To assess whether CITED2 overexpression also affected *in vivo* engraftment of human cells, $1\text{-}2x10^5$ transduced $CD34^+$ cells were transplanted into NSG mice (n=6/7) and at the indicated timepoints contribution of human $CD45^+$ cells to the peripheral blood was measured. Figure 1F and G demonstrate that $CD34^+$ cells overexpressing CITED2 contributed significantly better to human engraftment than control cells in 4 out of 7 transplanted mice ($p<0.05$). The higher engraftment in 4 mice suggested that CITED2 either affected migration/homing towards the BM or better maintained primitive HSCs. CXCR4 expression on both mRNA and protein level was not affected by CITED2 (*Supplementary Figures 1E* and *F*) and also the migration towards SDF-1 was not affected (*Supplementary Figure 1G*). However, within the bone marrow of the high engrafting mice, very primitive $lin^-CD34^+CD38^-CD90^+CD45RA^-$ cells could be detected at 28 weeks after transplantation, with a similar distribution of HSCs, MPPs and LMPPs as previously published[27] (*Supplementary Figure 1H*). These CITED2-maintained HSCs contributed normally to lineage differentiation, as CD33 myeloid, CD19 B-cells and CD3 T-cells were observed 24-27 weeks after transplantation (Figure 1H and *Supplementary figure 1I*).

**Enhanced CITED2 expression increases quiescence of human HSCs**

As CITED2 significantly improved the long-term *in vitro* and *in vivo* output, it suggested that CITED2 has a profound impact on more immature human HSCs. LTC-IC assays with the total $CD34^+$ compartment confirmed that CITED2 overexpression increases the LTC-IC frequency (*Supplementary figure 2A*). This coincided with a threefold increase in the percentage of immature $CD34^+CD38^-$ HSCs after 3 days of culture (Figures 2A and B). To gain further insight into how CITED2 increases the number of human HSCs, we analyzed whether these cells have a changed apoptotic or cell cycle profile. Figure 2C demonstrates that

***Figure 1. CITED2 enhances human HSPC cultures and engraftment***

***A**) Expansion on MS5 stromal co-culture of CB $CD34^+$ cells lentivirally transduced with CITED2 or control lentivirus. (n=4). Each week cultures were demi-populated, fresh media was added and cells were counted. The average of 4 independent experiments is shown. Error bars denote standard deviation. * $p < 0.05$. **B**) Image of typical cobblestone forming areas (at 10x magnification) after 5 weeks of culture on MS5 stromal cells. **C**) Colony forming cell numbers from $CD34^+$ cells directly after transduction with lentiviral CITED2 (n=8, error bars denote standard deviation * $p < 0.05$). **D**) Colony forming cell*

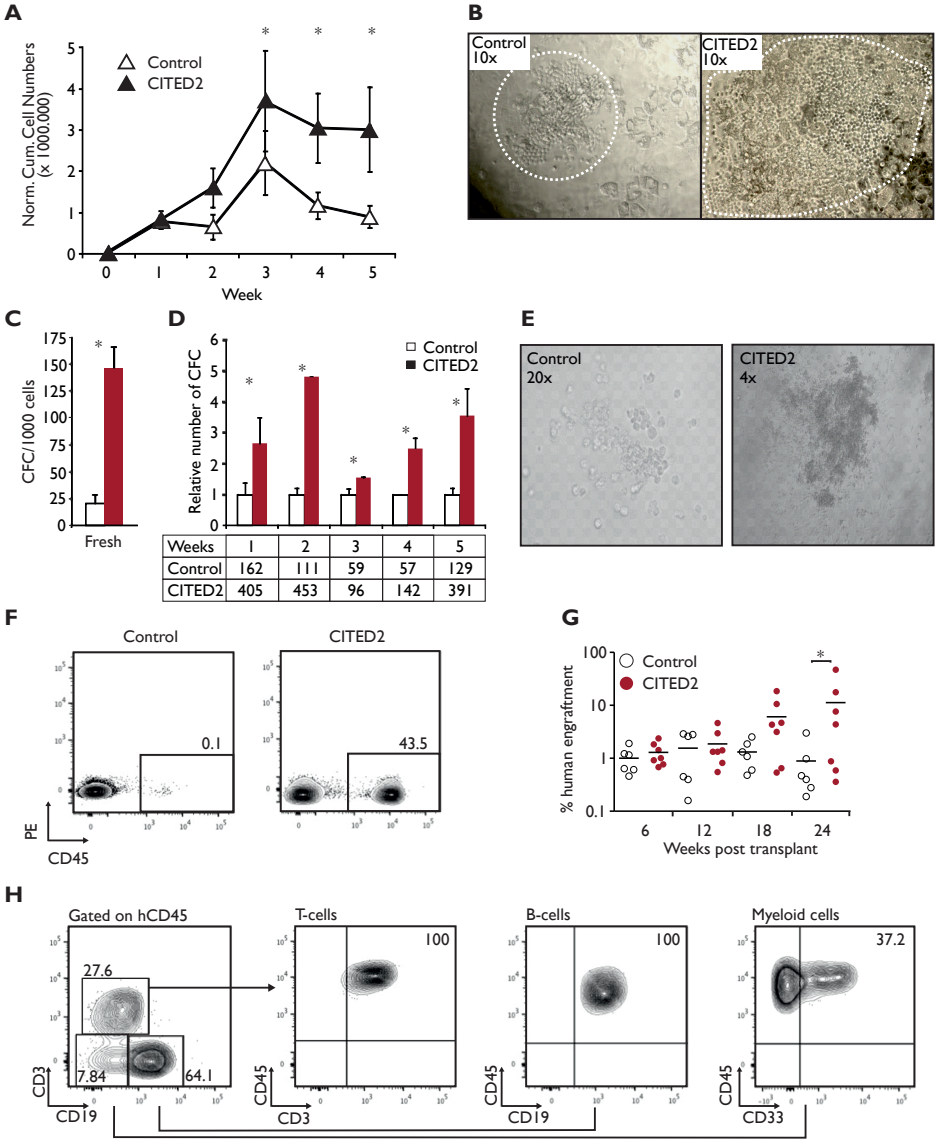



***Figure 1. CITED2 enhances human HSPC cultures and engraftment***

*numbers after the indicated weeks on MS5 stromal cells. Each week suspension cells were harvested and interrogated for CFC activity (n=3, error bars denote standard deviation). Relative CFC numbers are depicted for comparison, with average CFC numbers per 10.000 cells plated given underneath the figure. * $p < 0.05$. **E**) Typical images of control and CITED2-overexpressing colonies. Note that the control colonies have been taken at a 20x magnification, whereas the CITED2 transduced colonies have been taken at a 4x magnification. **F**) Typical FACS plots showing human engraftment in NSG mice 24 weeks post transplant with control or CITED2 transduced CD34$^+$ cells. **G**) Human engraftment in NSG mice transplanted with control (n=6) or CITED2 (n=7) transduced CD34+ cells. * $p < 0.05$. **H**) Typical FACS plots showing multilineage CD3 (T-cell), CD19 (B-cell) and CD33 (myeloid cell) engraftment at 24 weeks post transplant in a CITED2-engrafted mouse.*

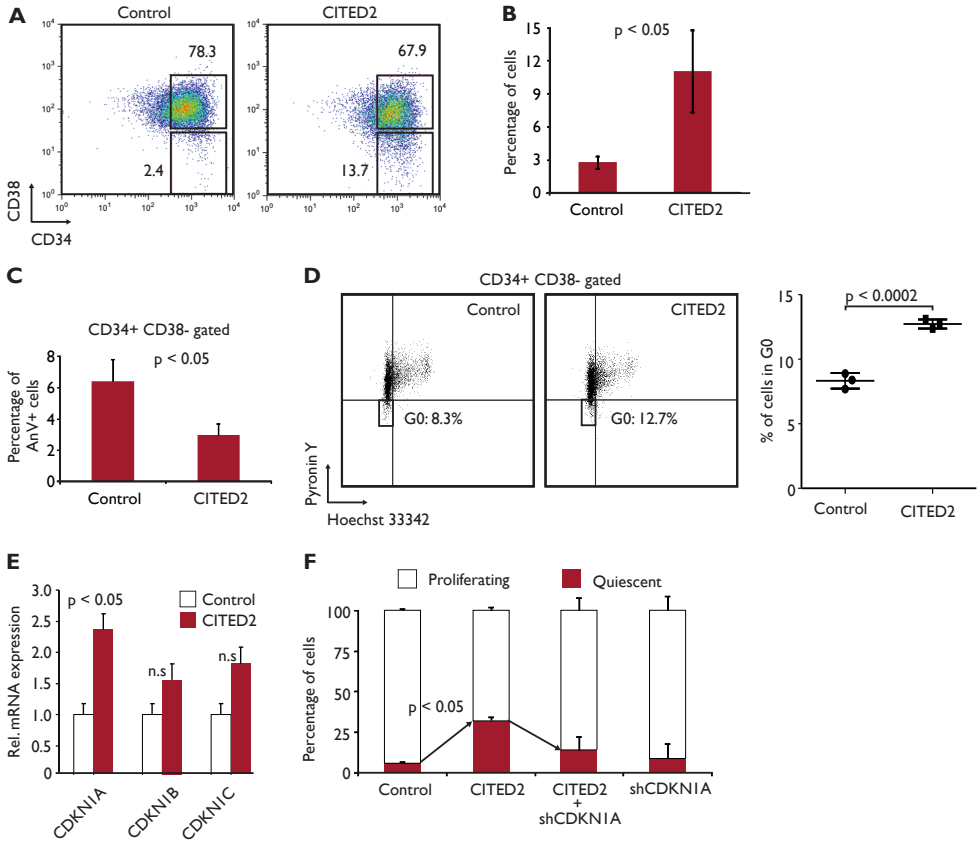



***Figure 2. Enhanced CITED2 expression increases quiescence of human hematopoietic stem cells***

***A**) CD34⁺ CB cells were lentivirally infected with control or CITED2 and cultured in HPGM with 100 ng/ml of SCF, TPO and FLT3. After 2-3 days cultures were analyzed for expression of CD34 and CD38. A typical FACS analysis of HSC (CD34⁺CD38⁻) or progenitor cells (CD34⁺CD38⁺ ) after transduction is shown (n=15). Cells were first gated through CD271 (tNGFR, transduced cells) and dapi (viability). **B**) Average percentages of CD34⁺CD38⁻ HSCs of 15 CBs. Error bars denote standard deviation. **C**) CB cells were transduced, cultured and stained for CD34, CD38, CD271 and AnnexinV. The average of 5 CBs is shown, with error bars denoting standard deviation. **D**) Cell cycle FACS analysis of transduced CD34⁺CD38⁻ cells. A representative example is shown. In the right scatter plot, the average and standard deviation of 3 samples is shown. **E**) Q-PCR analysis for CDKN1A, CDKN1B and CDKN1C after overexpression of CITED2 in CD34⁺CD38⁻ HSCs. **F**) CB CD34⁺ cells were lentivirally infected with control, CITED2 and a lentiviral short hairpin against CDKN1A. 120 single CD34⁺CD38⁻ HSCs were subsequently sorted into terasaki plates in IMDM plus 20% FCS and 10 ng/ml IL-3 and SCF. For 4 days of culture, each well was microscopically analyzed for cells that had divided or cells that had not divided (n=3). Individual experiments were normalized to compare. Error bars depict standard deviation.*

CITED2 decreases the percentage of CD34⁺CD38⁻ cells that are AnnexinV⁺. As murine HSCs cycle faster upon CITED2 deletion,[28] we assumed that overexpression of CITED2 might also affect human HSCs quiescence or proliferation. To confirm this, we performed Hoechst 33342/ PyroninY stainings to measure G0, G1 and G2/ S/M phases of the cell cycle. As shown in Figure 2D, an increased fraction of CITED2- transduced CD34⁺CD38⁻

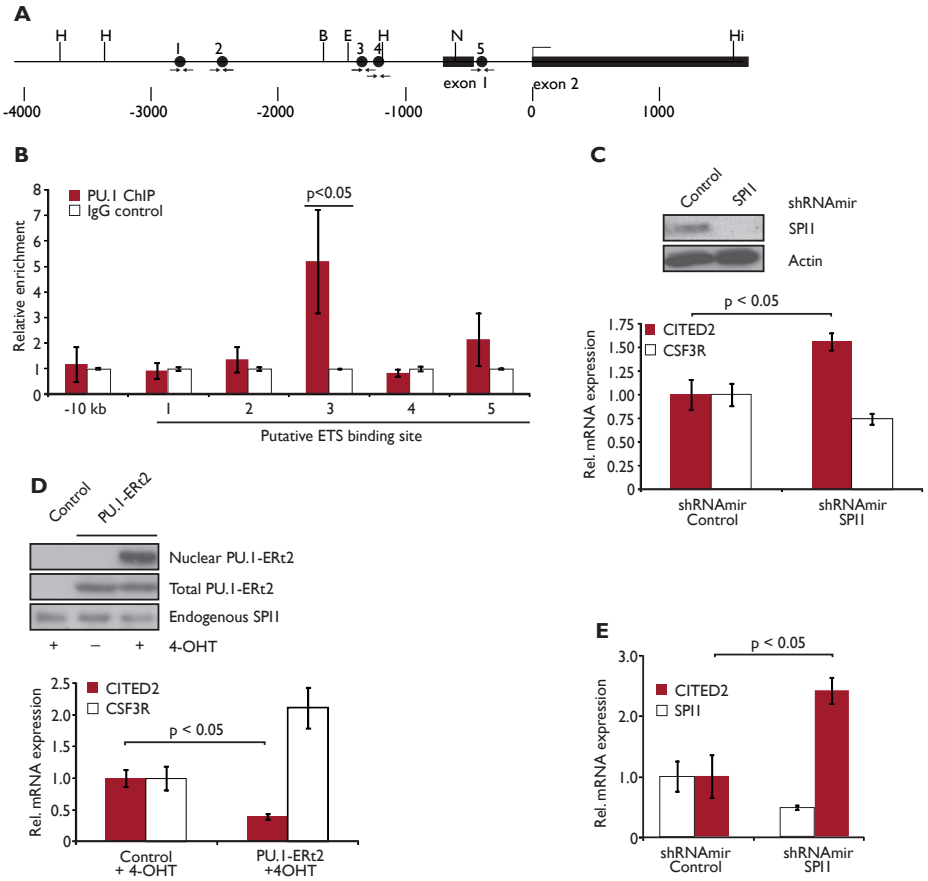



***Figure 3. CITED2 expression is repressed by PU.1***

***A**) Schematic representation of the human genomic CITED2 locus. Black bars indicate the two exons. The arrow in exon 2 indicates the start of the CITED2 CDS. Black dots indicate the positions of the 5 putative ETS binding sites (GAGGAA), with arrows indicating Chromatin Immuno Precipitation (ChIP) primers. B=BamHI, E=EcoRI, H=HinDIII, Hi=HincII, and N=NotI are indicated for reference. The bottom numbers indicate numbers of nucleotides away from the start ATG. **B**) Chromatin Immuno Precipitation (ChIP) for PU.1 on CB CD34⁺ cells. 2-5 million CD34⁺ cells were crosslinked for 15 min and ChIP was performed with anti-PU.1 ab (Santa Cruz, sc-352) or control pre-immune serum (IgG control). Recovered genomic DNA was subjected to Q-PCR with the indicated primers specific for the five putative ETS binding sites. A region 10 kb upstream was taken along as negative control. Data is presented as relative enrichment over IgG control, and is the average of 4 independent experiments, error bars denoting standard deviation. **C**) HL-60 cells were transduced with a lentivirus carrying a shRNAmir targeting human SPI1. Western blot analysis indicates proper knockdown (upper panel). mRNA was isolated and Q-PCR against CSF3R and CITED2 was performed (lower panel). **D**) HL-60 cells were transduced with lentiviral PU.1-ERt2 construct. After 4-OHT (100 nM) treatment for 4 days, nuclear lysates were isolated and analyzed by western blot (upper panel). Bottom panel: mRNA was isolated and Q-PCR was performed for CSF3R mRNA (to indicate proper PU.1 activation) and CITED2 mRNA. **E**) Primary CB-derived CD34⁺ cells were transduced with SPI1 shRNAmir and subjected to mRNA isolation and Q-PCR analysis for SPI1 and CITED2.*

HSCs remained in the G0 phase of the cell cycle. Q-PCR analysis demonstrated that increased CITED2 expression led to an increased expression of the cyclin-dependent kinase inhibitors CDKN1B, CDKN1C and especially CDKN1A in the

$CD34^+CD38^-$ compartment (Figure 2E), which is consistent with the enhanced quiescence of these cells. To asses directly the involvement of CDKN1A in the CITED2-mediated quiescence of HSCs, and to verify this observation in an independent assay, $CD34^+$ cells were doubly transduced with a lentivirus expressing CITED2 and a lentivirus-expressing a short hairpin against CDKN1A (*Supplementary figures 2B* and *C*). Subsequently, $CD34^+CD38^-$ HSCs were single cell sorted and their proliferation was microscopically evaluated. For 4 days, each well was inspected for the presence of one cell (quiescence) or more than one cell (proliferation). CITED2 expression indeed induced more cells to remain quiescent (Figure 2F). Furthermore, this experiment demonstrated that the CITED2-induced quiescence is partially reversible on knockdown of CDKN1A. This indicates that CITED2 can modulate the quiescence of HSCs by regulating the expression of CDKN1A and thereby maintain HSCs. Although this CITED2-induced HSC quiescence is consistent with the enhanced long-term engraftment *in vivo*, it seems contradictory to the observed expansion in our *in vitro* assays. However, performing the single-cell quiescence/proliferation assay also on CITED2-transduced $CD34^+CD38^+$ progenitor cells demonstrated that $CD34^+CD38^+$ progenitors are actually induced to proliferate (*Supplementary figure 2D*) upon transduction with CITED2, demonstrating differential effects of CITED2 on the HSC and progenitor compartments.

**CITED2 is a direct repressive target of PU.1**

Next, we investigated the mechanisms underlying the enhanced CITED2 expression. Besides the well studied HIF1α,[24] FOXO3A[9] and STAT5[10] binding sites, the human CITED2 promoter containes five potential PU.1-binding sites (Figure 3A), suggesting that PU.1 has a role in regulating CITED2 expression. ENCODE[29] ChIP-seq data for PU.1 in hematopoietic cell lines confirmed this hypothesis (*Supplementary figure 3A*). In addition, we first assessed PU.1 binding to the CITED2 promoter under relevant conditions, by performing chromatin immunoprecipitations (ChIPs) for endogenous PU.1 in CB-derived $CD34^+$ cells. This demonstrated that endogenous PU.1 indeed binds to the CITED2 promoter (Figure 3B). PU.1 binding to the CITED2 promoter was independently verfied in PU.1-overexpressing 293T lysates with streptavidin pulldown assays. Biotinylated oligos containing the PU.1-binding sites from the CITED2 promoter bound PU.1 as efficiently as a bona fide PU.1-binding site from the JUNB promoter (*Supplementary figure 3C*). Subsequently, *SPI1* (the gene encoding human PU.1) was downregulated by means of a short hairpin Mir-based lentivirus (Figure 3C and *Supplementary figure 3D*). This resulted in an increase in CITED2 expression in both HL-60 and CB $CD34^+$ cells (Figures 3C and E). In contrast, overexpression of a 4-hydrotamoxifen (4-OHT)-inducible PU.1 (PU.1-ERt2; Figure 3D and *Supplementary figure 3E*) in HL-60 cells led to ~2-fold decrease in CITED2 expression (Figure 3D).

To study the molecular interaction between PU.1 and CITED2 in more detail, the human kidney cell line 293T was used. This cell line expresses high levels of CITED2 and no PU.1. 293T cells were transduced with PU.1-ERt2 and activation of PU.1 resulted in expression of its target genes NCF4 (*Supplementary figure 3B*) and in a twofold reduction in CITED2 expression, validating this model for further studies. To prove that CITED2 is indeed a direct target of PU.1, the CITED2

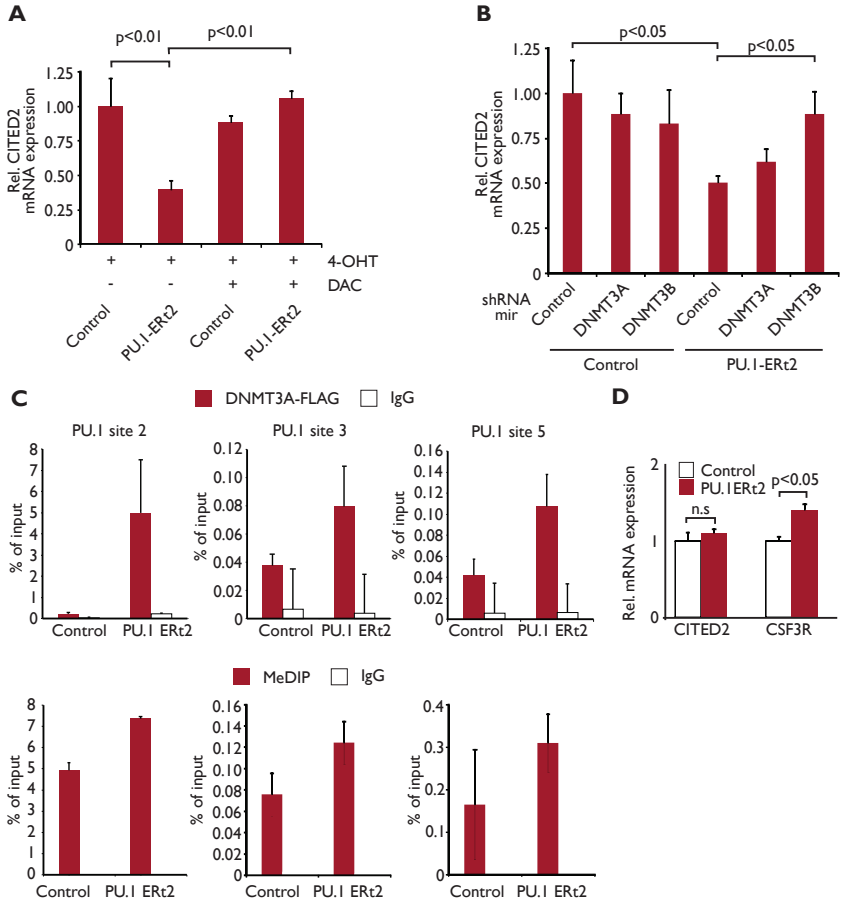



***Figure 4. PU.1 represses CITED2 expression through DNMT3A and B***

***A**) CB CD34⁺ cells were transduced with PU.1-ERt2 or control vector and stimulated with 100 nM 4-OH tamoxifen up to 96 hrs, with or without 0.5 M Decitabine (DAC). Q-PCR analysis was performed to investigate CITED2 mRNA expression. **B**) 293T cells infected with PU.1-ERt2 or control vector were transduced with pGIPZ SFFV shRNAmir Scrambled (SCR), DNMT3A or DNMT3B and stimulated with 100 nM 4-OH tamoxifen up to 96 hrs. Data is normalized to controls. QPCR analysis was performed to investigate CITED2 mRNA expression (n=3). **C**) 293T cells infected with PU.1-ERt2 or control vector were transduced with a FLAG-tagged DNMT3A construct and stimulated with 100 nM 4-OH tamoxifen up to 96 hrs. ChIP for DNMT3A ChIP was performed with anti-FLAG (M2), an antibody that specifically recognizes methylated DNA (MeDIP) or control pre-immune serum (IgG control). Recovered genomic DNA was subjected to Q-PCR with the indicated primers specific for the 3 PU.1 binding sites that showed the strongest PU.1 binding. **D**) The OCI-AML3 cell-line, harboring a DNMT3AR882C mutation, was transduced with lentiviral PU.1-ERt2, and stimulated with 100 nM 4-OHT for 4 days. mRNA was isolated to investigate CITED2 expression (n=3).*

promoter was subcloned upstream of a luciferase and/or GFP construct. After mutating the strongest ChIP-binding sites (sites 3 and 5), reporter activity was assessed (*Supplementary figure 3F*). Supplementary Figure 3G shows that activation of PU.1 reduces activity of the wild-type CITED2 (pCT2-wt) reporter. This reporter thus responds in the same manner as the endogenous CITED2

promoter, validating it for further experiments. Mutation of binding site 3 showed a moderate effect (*Supplementary figures 3H* and *I*), but mutation of site 5 (pCT2-Δ5 and pCT2-Δ3Δ5) rescued the repression of CITED2 by PU.1 to a large extent (*Supplementary figures 3H* and *I*). Taken together, these data demonstrate that CITED2 is a direct target of PU.1 and that activation of PU.1 leads to repression of CITED2.

**PU.1 represses CITED2 expression through DNMT3A and DNMT3B**

As PU.1 has been shown to mediate target gene repression via DNA methylation,[30] we investigated whether methylation was also involved in repression of CITED2. Therefore, CB CD34$^+$ cells were transduced with PU.1-ERt2 and treated for 4 days with 100 nM 4-OHT in the presence or absence of 0.5 μM decitabine, a DNA methyltransferase inhibitor. Decitabine completely abrogated PU.1-mediated repression of CITED2 (Figure 4A), suggesting that PU.1 recruits DNA methyltransferases (DNMTs) to repress CITED2 expression. As CITED2 expression is highest in immature cells and gradually decreases during myelopoiesis,[2] the *de novo* methylases DNMT3A or DNMT3B are potentially involved. To this end, PU.1-ERt2-expressing 293Ts were transduced with lentiviruses encoding short hairpins targeting either DNMT3A or DNMT3B (*Supplementary figures 4A–C*). The scrambled control-infected cells showed repression of CITED2 expression upon activation of PU.1 (Figure 4B, control groups). In contrast, suppressing DNMT3A or DNMT3B by means of RNA interference rescued the repression of the CITED2 promoter by PU.1 either partially (DNMT3A) or almost completely (DNMT3B) (Figure 4B).

Both DNMT3A and DNMT3B function synergistically in the same complex,[31] but DNMT3A is the most frequently mutated DNMT3 gene in AML.[32] As deletion of Dnmt3a in mice also resulted in expansion of HSC numbers,[33] we further focused on DNMT3A. Next, we cotransduced PU.1-ERt2-expressing 293Ts with lentiviruses encoding a FLAG-tagged DNMT3A construct and treated for 4 days with 100 nM 4-OHT. A subsequent ChIP against FLAG demonstrated that activation of PU.1 led indeed to recruitment of DNMT3A to the CITED2 promoter at the sites that showed the strongest PU.1 binding (Figure 4C, top panels). This PU.1-mediated recruitment of DNMT3A coincided with a small increase in DNA methylation as measured through ChIP for methylated DNA (Figure 4C, bottom panels). Based on these data we hypothesized that in cells bearing mutations in DNMT3A, PU.1 should no longer be able to repress CITED2 expression (or to a lesser extent). To test this, the OCI-AML3 cell line with a DNMT3A$^{R882C}$ mutation was transduced with PU.1-ERt2 and stimulated with 4-OHT for 4 days. Indeed, PU.1 did not repress CITED2 expression in this cell line (Figure 4D), although PU.1 was activated as measured through CSF3R expression. Taken together, these data show that DNMT3A and DNMT3B have nonredundant functions in the PU.1-mediated repression of CITED2.

**CITED2 and SPI1 expression are inversely correlated in leukemic CD34+ cells**

Myeloid transcription factors such as C/EBPα and PU.1 are frequently inactivated in AML.[12,23] Having established that PU.1 represses CITED2 expression, we investigated whether CITED2 expression was enhanced in AML. CITED2 mRNA expression was analyzed by Q-PCR in AML CD34$^+$ cells (n=28) and compared with CITED2 expression in CD34$^+$ cells from normal

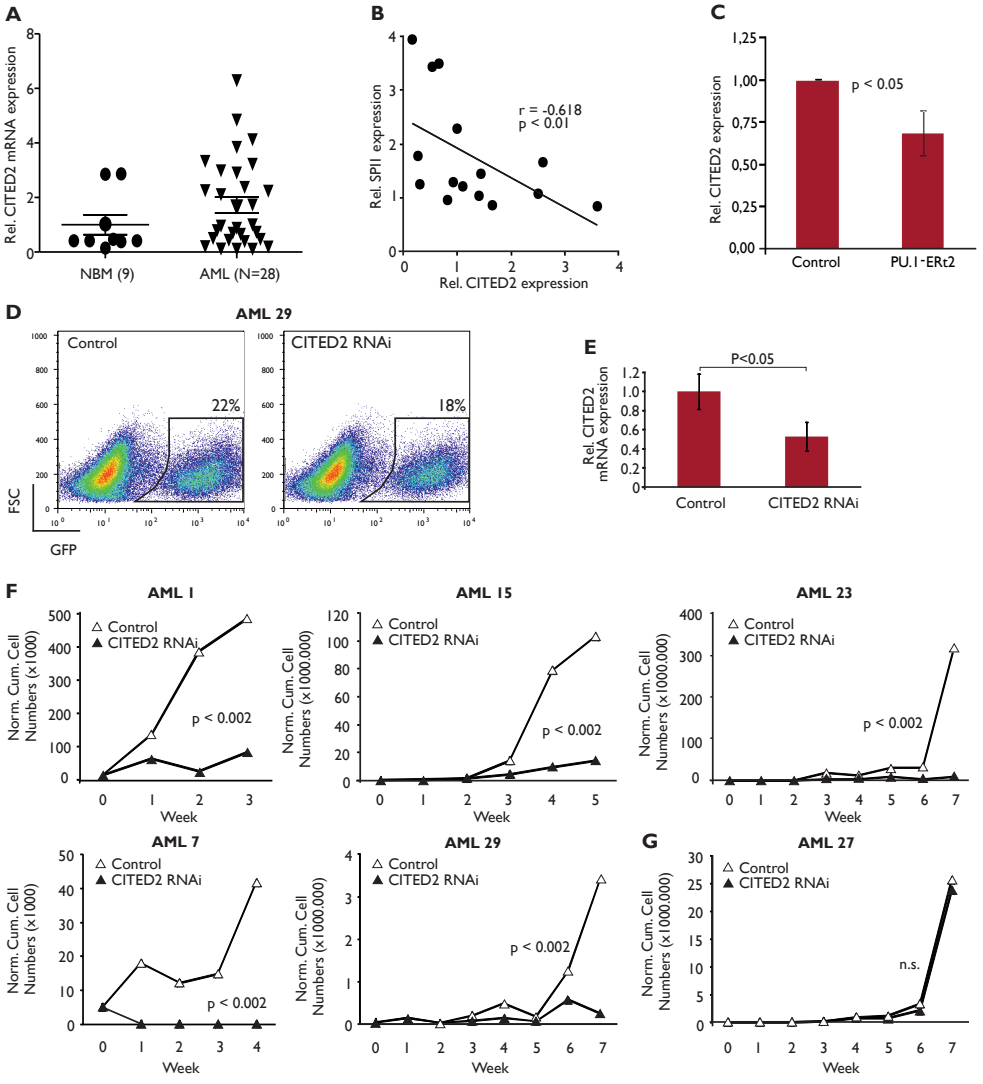



***Figure 5. CITED2 expression is necessary for leukemic stem cell maintenance***

***A**) CD34⁺ cells from acute myeloid leukemia (AML, n=28) and CD34⁺ cells from normal bone marrow (NBM, n=9) were isolated and investigated for CITED2 mRNA expression by Q-PCR. **B**) Scatter plot analysis of CITED2 expression within CD34⁺ cells from AMLs with varying degrees of SPI1 expression. The spearman's ranked correlation coefficient is indicated. **C**) CD34⁺ cells from AML no. 1 and 29 were transduced with lentiviral PU.1-ERt2 or control, 4-OHT treated for 3 days and mRNA was isolated. Subsequently Q-PCR was performed for CITED2. **D**) Representative example of AML CD34⁺ cells transduced with control hairpins (scrambled control) or short hairpins against CITED2, indicating similar transduction efficiencies. **E**) Sorting lentivirally transduced CD34⁺ cells from AML patients (n=7) demonstrates a ~55% knockdown of CITED2 mRNA expression, as measured by Q-PCR. Mean of the AMLs is presented with error bars denoting standard deviation. **F**) Lentiviral knockdown of CITED2 in CD34⁺ AML cells demonstrate that CITED2 expression in these cells is essential for inititation of long-term leukemic cultures on MS5 stromal layers. Five representative growth curves are shown. **G**) One representative AML from the nonresponsive group with low CITED2 expression demonstrates that CITED2 is not essential for the initiation of leukemic growth in this particular AML.*

*Table 1. Patient characteristics*

| AML Patient | FAB | Risk | FLT3 | NPM | DNMT3A | Cytogenetics | CITED2 knock-down | PU.1 overexpression | % transduction Control | % transduction shCITED2 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | M2 | Good | Wt | n.d. | Wt | t(8.21) | Growth impaired | CITED2 expression down | 45 | 35 |
| 2 | M1 | Intermediate | Wt | n.d. | Wt | NK | | | | |
| 3 | M5 | Poor | Wt | n.d. | Wt | 5q-;-7;2p-;17p | | | | |
| 4 | M5 | Unknown | ITD | Wt | Wt | n.d. | | | | |
| 5 | M4 | Intermediate | Wt | n.d. | Wt | NK | | | | |
| 6 | M4 | Intermediate | Wt | Wt | Wt | inv16 | Failed expansion | | 21 | 17 |
| 7 | M1 | Intermediate | n.d. | Wt | R882C | NK | Growth impaired | no downregulation of CITED2 | 19 | 17 |
| 8 | M1 | Poor | Wt | n.d. | Wt | inv.(3q),7-,10- | | | | |
| 9 | M4 | Good | Wt | Wt | n.d. | Inv16,c-kit+ | | | | |
| 10 | n.d. | Intermediate | Wt | Wt | R882H | NK | Failed expansion | no downregulation of CITED2 | 50 | 54 |
| 11 | M1 | Intermediate | ITD | Cyt | R882C | NK | Growth impaired | no downregulation of CITED2 | 19 | 17 |
| 12 | M4eo | Good | Wt | Wt | n.d. | t(16;16)(p13;q22), c-kit+ | | | | |
| 13 | M1 | Poor | ITD | Wt | Wt | NK | | | | |
| 14 | M5 | Intermediate | Wt | n.d. | Wt | NK | | | | |
| 15 | M0 | Intermediate | ITD | Wt | Wt | NK | Growth impaired | | 12 | 13 |
| 16 | M5 | Intermediate | ITD | n.d. | Wt | NK | | | | |
| 17 | M4 | Intermediate | Wt | n.d. | Wt | NK | | | | |
| 18 | M1 | Unknown | ITD | n.d. | Wt | NK | | | | |
| 19 | M5 | Intermediate | ITD | Wt | Wt | NK | Growth impaired | | 27 | 23 |
| 20 | M0 | Poor | n.d. | Wt | Wt | 5q-,Trisomy 6 | Failed expansion | | 14 | 16 |
| 21 | M2 | Intermediate | ITD | Cyt | R882H | NK | | no downregulation of CITED2 | | |
| 22 | M2 | Intermediate | ITD | Wt | Wt | NK | No response | | 22 | 19 |
| 23 | M5 | Intermediate | ITD | Cyt | Wt | NK | Growth impaired | | 28 | 24 |
| 24 | M5 | Unknown | Wt | Wt | Wt | NK | Growth impaired | | 48 | 48 |
| 25 | M2 | Poor | Wt | Wt | Wt | 45,X,-Y,t(8;21),13q-,14q+[8]/46,XY[2] | Growth impaired | | 22 | 21 |
| 26 | M2 | Intermediate | n.d. | Wt | Wt | t(8;21),t(q22;q22) | | | | |
| 27 | M1 | Poor | Wt | Wt | n.d. | 45,XX, 3p+, -7, 8p-, 46,XX | No response | | 10 | 10 |
| 28 | M1 | Poor | ITD | Wt | Wt | NK | | | | |
| 29 | M2 | Poor | Wt | Cyt | Wt | -7, -10 | Growth impaired | CITED2 expression down | 22 | 22 |
| 30 | M1 | Poor | Wt | | n.d. | inv.(3q),7-,10- | Growth impaired in vivo | | 3 | 5 |
| 31 | M2 | Poor | ITD | Wt | n.d. | 47,XY,t(3;5)(q23;q33),+8[10] | Growth impaired in vivo | | 8 | 12 |
| 32 | n.d. | Poor | ITD | Cyt | n.d. | complex | Growth impaired in vivo | | 13 | 19 |

*AML no 29 gave consistent poor-quality RNA, so no mRNA expression analysis could be performed

n.d. = not determined

bone marrow (n=9). Figure 5A shows a wide range of CITED2 expression in AML CD34$^+$ cells, as compared with normal bone marrow CD34$^+$. Thirteen out of 28 AML patients displayed higher than normal expression. Similar results were obtained when compared with CD34$^+$ cells from CB (n=4, *data not shown*) or granulocyte-colony stimulating factor-mobilized peripheral blood stem cells (n=5, *data not shown*). No apparent correlation between CITED2 mRNA expresssion and patient characteristics could be observed in our data set (Table 1). To verify these results, we analyzed two independent data sets. First, the difference in CITED2 expression between normal human bone marrow subsets and CD34$^+$ selected AML samples with defined translocations (n=142) was investigated using the HemaExplorer website (*http://servers.binf.ku.dk/hema-explorer/*).[34] *Supplementary Figure 5A* demonstrates that CITED2 expression was significantly higher in most AML samples as compared with nomal cellular bone marrow subsets. Second, also in CD34$^+$ cells from pediatric AML samples (n=17, profiled by Andersson *et al.*[6]) CITED2 expression was found to be significantly higher as compared with CD34$^+$ normal bone marrow cells (*Supplementary Figure 5B*). Taken together, these data indicate that CD34$^+$ cells from AML patients have higher CITED2 expression than normal CD34$^+$ cells. Next, we examined whether a negative correlation exists between PU.1 and CITED2 expression. We therefore plotted *SPI1* expression against CITED2 expression and we observed a significant inverse correlation between *SPI1* and CITED2 expression (Figure 5B and *Supplementary Figure 5C*). To exclude differentiation-dependent effects, we phenotyped a panel of these AML samples for a lymphoid-primed multipotent progenitor- or granulocyte–macrophage progenitor-like phenotype.[35] CD45RA was strongly expressed in all AMLs analyzed, in both CD38$^-$ and CD38$^+$ compartments, suggesting that all AMLs displayed a granulocyte–macrophage progenitorlike mature phenotype (*Supplementary Figure 5D*). As CITED2 expression is low in normal granulocyte–macrophage progenitors (*Supplementary Figure 5A* and *data not shown*), this supported our finding that CITED2 expression is aberrantly high in these AMLs. Subsequent bisulfate conversion and pyrosequencing of 12 AML samples demonstrated a general hypomethylation of the CITED2 promoter around the PU.1-binding sites (*Supplementary Figure 5E*), consistent with the higher expression of CITED2.

Next, we investigated whether PU.1 also regulates CITED2 expression in patient AML cells. AML CD34$^+$ cells were transduced with PU.1-ERt2. AML numbers 1 and 29 responded with a decreased expression of CITED2 upon 4-OHT treatment (Figure 5C and Table 1). In contrast, 4 of the 29 AML patient samples (AML numbers 7, 10, 11 and 21) contained a DNMT3A$^{R882C/H}$ mutation (Table 1), and did not show a decrease in CITED2 expression upon PU.1-ERt2 activation (Table 1), similar to the results in the OCI-AML3 cell line (Figure 5D). This supports our data that DNMT3A is critical for PU.1 to repress CITED2, and together these data indicate that loss of PU.1 or DNMT3A is a contributing factor in elevating CITED2 levels.

### CITED2 is required for maintenance of leukemic cultures

Gene set enrichment analysis on gene expression profiles of CD34$^+$CD38$^-$ HSCs after CITED2 upregulation indicated that various leukemia-related signatures are enriched upon overexpression of CITED2 (*Supplementary Table 1*), suggesting that CITED2 expression has a functional role in AML maintenance. We performed RNA interference against

CITED2 in CD34$^+$ AML cells (n = 14). AML CD34$^+$ cells, transduced with lentiviral RNA interference vectors targeting CITED2 or a scrambled hairpin (control) (Figures 5D and E), were cultured on MS5 stromal cells, as described previously,[36] and observed for several weeks. Two different hairpins were used to knockdown CITED2, to minimize off target effects, both with similar efficiencies and cell biologic results. (*Supplementary Figures 6A* and *D*, *Supplementary Table 2*). Control cells from AML numbers 6, 10 and 20 failed to expand on MS5 stromal cultures and were not analyzed further. Eighty-two percent of the AMLs (9/11) showed a severely impaired expansion upon knockdown of CITED2 (Figure 5F and Table 1). This most likely is the result of combined effects on cell cycle as well as on apoptosis. We have also analyzed the differentiation of these AMLs *in vitro*, but did not observe any differences between shCITED2- and control-transduced cells (n = 11, *data not shown*). From several AMLs we compared the level of knockdown to the relative decrease in leukemic expansion, but we could not detect a significant correlation between the two (*Supplementary Fig. 6E*). Lastly, we addressed whether knockdown of CITED2 also inhibits AML engraftment in vivo. We transduced CD34$^+$ AML cells from AML numbers 30–32 with lentiviral control or CITED2 RNAi and transplanted 40, 25 or 18% GFP-sorted cells into NSG mice (N = 17). At 9 weeks post-transplant the peripheral blood of these mice was analyzed for the presence of CD45$^+$GFP$^+$ cells. *Supplementary Figures 6F* and *G* demonstrate that the mice that received CITED2-knockdown AML cells had a significant lower contribution of GFP$^+$ cells, as compared to the mice that received control transduced AML cells. As the phenotype of the transplanted AML cells was CD34$^+$CD33$^+$ (*Supplementary Figure 6G*), these cells represent cells of leukemic origin. Taken together, these data demonstrate that CITED2 is highly expressed in a subset of AML CD34$^+$ cells, and that CITED2 has a critical function in maintaining long-term leukemic growth both *in vitro* and *in vivo*.

## Discussion

We present data showing that enhanced expression of CITED2 influences the function of normal human HSPCs *in vitro* and *in vivo* and that interfering with CITED2 expression has a profound impact on primary human leukemic cells. We clearly demonstrate that CITED2 expression enhances *in vitro* long-term culture and *in vivo* engraftment into NSG mice. Our human data are therefore not only consistent with observations in mice that CITED2 is required for proper murine HSCs function, but additionally demonstrate that overexpression of CITED2 alone is sufficient to maintain the primitive CD34$^+$CD38$^-$ HSC pool by decreasing apoptosis and enhancing quiescence. This is in line with data from CITED2$^{\Delta/\Delta}$ mice, where deletion of CITED2 resulted in actively cycling long-term HSCs with an increased apoptotic profile.[2,28] Our data indicate that in human HSCs CITED2 induces the expression of CDKN1A and CDKN1C. This is slightly different from the experiments performed by Du et al.,[28] as they have shown that CITED2 controls the expression of the cyclin-dependent kinase inhibitor Cdkn1c, but not of Cdkn1a.[28] The mechanism behind this is at present unclear. CITED2 probably does not bind to DNA itself, but most likely uses CBP/p300 in response to various external signals.[1,5] We functionally validated these data demonstrating that downregulation of CDKN1A can rescue the CITED2-induced quiescence in CD34$^+$CD38$^-$ HSCs. Deletion of Cdkn1a or Cdkn1c in mice resulted in a loss of engrafting

cells upon serial transplantation.[37,38] Thus by inducing CDKN1A/C-mediated quiescence, CITED2 prevents exhaustion and hence increases overall hematopoiesis, as is demonstrated by our enhanced long-term culture and *in vivo* engraftment data.

It appears that PU.1, one of the main transcription factors involved in HSC maintenance and differentiation, affects CITED2 expression. Lowering PU.1 activation in human $CD34^+$ cells led to an increase in CITED2 expression, and conversely, activating PU.1 led to a decrease in CITED2 expression. Previously, we observed that HSCs from the PU.1-knockout mice also showed a fourfold increase in the expression of CITED2 (*data not shown*) and maintained their long-term HSCs. This therefore suggested that CITED2 contributes to the maintenance of these long-term HSCs and the subsequent development of AML.[18,19,39]

In AML, such a direct inverse correlation between CITED2 and SPI1 expression was observed in many, but not all cases (Figure 5B). This most likely stems from the fact that many oncogenes affect the activation of PU.1 rather than its expression,[11,13,17] as well as the interplay with various cytokines and transcription factors that activate CITED2 expression.[5,9,10] Nevertheless, the decreased expression of CITED2 upon PU.1 reactivation in these AMLs is in line with our other data.

In addition, our data demonstrated that both DNMT3A and DNMT3B knockdown interfered with PU.1-mediated repression of CITED2, similar to its repressive function on p16INK4A.[30] Since in AML both DNMT3A and DNMT3B mutations have been observed[32,40] and both factors are necessary for forming repressive complexes,[31] our data are consistent with such observations. Activation of PU.1 led to a ~2- to ~10-fold enrichment of DNMT3A binding around the PU.1 binding sites on the CITED2 promoter (Figure 4C), although because of technical variation in ChIP efficiency this was not significant. A similar trend was observed toward an increase in methylation upon PU.1 activation (Figure 4C), as could be expected from a de novo methyltransferase. Bisulfate conversion and pyrosequencing in 12 of our AML samples demonstrated a general hypomethylation of the CITED2 promoter around the PU.1 binding sites (*Supplementary Figure 5E*), which was consistent with the overall enhanced CITED2 expression in AML. However, it must be noted that the CITED2 promoter is a large CpG island, which are typically hypomethylated.[41,42] Furthermore, we cannot exclude the possibility that PU.1 and DNMT3A affect the methylation of a currently unknown upstream enhancer of CITED2. Our observations that in AMLs with a DNMT3A mutation (discovered in ~20% of all AML patients[32]) PU.1 is unable to repress CITED2 expression is consistent with such a possibility.

Whether CITED2 only contributes to AML maintenance downstream of PU.1 or is also involved in leukemia initiation is not resolved. Our data functionally demonstrate that CITED2 is essential for long-term leukemic maintenance *in vitro* and *in vivo* and our GSEA results indicate that the CITED2-induced gene expression programs at least partially overlap with published leukemia-related signatures (*Supplementary Table 1*).[43-47] Although overexpression of CITED2 changed proliferation and apoptosis of HSCs and skewed myeloerythroid differentiation of progenitors, no leukemic transformation was observed. This suggests that additional alterations are necessary.

Taken together, our findings indicate that CITED2 has a critical role in regulating normal versus malignant

hematopoiesis. We propose a model where CITED2 repression during myelopoiesis is among others, regulated by PU.1, via DNMT3A/B. Together with diminished activation of PU.1, this results in a perturbed myeloid differentiation program and is likely to contribute to leukemic development and/or maintenance.

## References


*1 Yoon H, Lim J-H, Cho C-H, et al. CITED2 controls the hypoxic signaling by snatching p300 from the two distinct activation domains of HIF-1α. Biochim Biophys Acta 2011.*

*2 Kranc KR, Schepers H, Rodrigues NP, et al. Cited2 is an essential regulator of adult hematopoietic stem cells. Cell Stem Cell 2009;*

*3 Kranc KR, Bamforth SD, Bragança J, et al. Transcriptional coactivator Cited2 induces Bmi1 and Mel18 and controls fibroblast proliferation via Ink4a/ARF. Mol Cell Biol 2003;*

*4 Berlato C, Chan KV, Price AM, et al. Alternative TFAP2A isoforms have distinct activities in breast cancer. Breast Cancer Res 2011;*

*5 Chou Y-T, Hsieh C-H, Chiou S-H, et al. CITED2 functions as a molecular switch of cytokine-induced proliferation and quiescence. Cell Death Differ 2012;*

*6 Andersson A, Ritz C, Lindgren D, et al. Microarray-based classification of a consecutive series of 121 childhood acute leukemias: prediction of leukemic and genetic subtype as well as of minimal residual disease status. Leukemia 2007;*

*7 Radich JP, Dai H, Mao M, et al. Gene expression changes associated with progression and response in chronic myeloid leukemia. Proc Natl Acad Sci U S A 2006;*

*8 Diaz-Blanco E, Bruns I, Neumann F, et al. Molecular signature of CD34(+) hematopoietic stem and progenitor cells of patients with CML in chronic phase. Leukemia 2007;*

*9 Bakker W, Harris I, Mak T. FOXO3a is activated in response to hypoxic stress and inhibits HIF1-induced apoptosis via regulation of CITED2. Mol Cell 2007;*

*10 Bakker W, van Dijk T, Parren-van Amelsvoort M, et al. Differential regulation of Foxo3a target genes in erythropoiesis. Mol Cell Biol 2007;*

*11 Cook W, McCaw B, Herring C, et al. PU.1 is a suppressor of myeloid leukemia, inactivated in mice by gene deletion and mutation of its DNA binding domain. Blood 2004;*

*12 Tenen D. Disruption of differentiation in human cancer: AML shows the way. Nat Rev Cancer 2003;*

*13 Vangala R, Heiss-Neumann M, Rangatia J, et al. The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. Blood 2003;*

*14 Laricchia-Robbio L, Premanand K, Rinaldi C, et al. EVI1 Impairs Myelopoiesis by Deregulation of PU.1 Function. Cancer Res 2009;*

*15 Walter M, Park J, Ries R, et al. Reduced PU.1 expression causes myeloid progenitor expansion and increased leukemia penetrance in mice expressing PML-RARalpha. Proc Natl Acad Sci U S A 2005;*

*16 Yang H, Liang H, Yan J-S, et al. Down-regulation of hematopoiesis master regulator PU.1 via aberrant methylation in chronicmyeloid leukemia. Int J Hematol 2012;*

*17 Mizuki M, Schwable J, Steur C, et al. Suppression of myeloid transcription factors and induction of STAT response genes by AML-specific Flt3 mutations. Blood 2003;*

*18 Rosenbauer F, Wagner K, Kutok J, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. Nat Genet 2004;*

*19 Metcalf D, Dakic A, Mifsud S, et al. Inactivation of PU.1 in adult mice leads to the development of myeloid leukemia. Proc Natl Acad Sci U S A 2006;*

*20 Houston IB, Kamath MB, Schweitzer BL, et al. Reduction in PU.1 activity results in a block to B-cell development, abnormal myeloid proliferation, and neonatal lethality. Exp Hematol 2007;*

*21 Aue G, Du Y, Cleveland SM, et al. Sox4 cooperates with PU.1 haploinsufficiency in murine myeloid leukemia. Blood 2011;*

*22 Schepers H, van Gosliga D, Wierenga ATJ, et al STAT5 is required for long-term maintenance of normal and leukemic human stem/progenitor cells. Blood 2007;*

*23 Schepers H, Wierenga ATJ, van Gosliga*

*D, et al. Reintroduction of C/EBPalpha in leukemic CD34+ stem/progenitor cells impairs self-renewal and partially restores myelopoiesis. Blood 2007;*

**24** *Leung MK, Jones T, Michels CL, et al. Molecular cloning and chromosomal localization of the human CITED2 gene encoding p35srj/Mrg1. Genomics 1999;*

**25** *Bhattacharya S, Michels C, Leung M, et al. Functional role of p35srj, a novel p300/CBP binding protein, during transactivation by HIF1. Genes Dev 1999;*

**26** *van den Boom V, Rozenveld-Geugien M, Bonardi F, et al. Nonredundant and locus-specific gene repression functions of PRC1 paralog family members in human hematopoietic stem/progenitor cells. Blood 2013;*

**27** *Majeti R, Park CY, Weissman IL. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell 2007;*

**28** *Du J, Chen Y, Li Q, et al. HIF-1α deletion partially rescues defects of hematopoietic stem cell quiescence caused by Cited2 deficiency. Blood 2012;*

**29** *Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome Res 2002;*

**30** *Suzuki M, Yamada T, Kihara-Negishi F, et al. Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. Oncogene 2006;*

**31** *Li J-Y, Pu M-T, Hirasawa R, et al. Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. Mol Cell Biol 2007;*

**32** *Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med 2010;*

**33** *Challen GA, Sun D, Jeong M, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. Nat Genet 2012;*

**34** *Bagger FO, Rapin N, Theilgaard-Mönch K, et al. HemaExplorer: a database of mRNA expression profiles in normal and malignant haematopoiesis. Nucleic Acids Res 2013;*

**35** *Goardon N, Marchi E, Atzberger A, et al. Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer Cell 2011;*

**36** *van Gosliga D, Schepers H, Rizo A, et al. Establishing long-term cultures with self-renewing acute myeloid leukemia stem/progenitor cells. Exp Hematol 2007;*

**37** *Matsumoto A, Takeishi S, Kanie T, et al. p57 is required for quiescence and maintenance of adult hematopoietic stem cells. Cell Stem Cell 2011;*

**38** *Cheng T, Rodrigues N, Shen H, et al. Hematopoietic stem cell quiescence maintained by p21cip1/waf1. Science 2000;*

**39** *Steidl U, Rosenbauer F, Verhaak RGW, et al. Essential role of Jun family transcription factors in PU.1 knockdown–induced leukemic stem cells. Nat Genet 2006;*

**40** *Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 2013;*

**41** *Galm O, Herman JG, Baylin SB. The fundamental role of epigenetics in hematopoietic malignancies. Blood Rev 2006;*

**42** *Bird A, Taggart M, Frommer M, et al fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell 1985;*

**43** *Valk P, Verhaak R, Beijen M, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med 2004;*

**44** *Graham SM, Vass JK, Holyoake TL, et al.. Transcriptional analysis of quiescent and proliferating CD34+ human hemopoietic cells from normal and chronic myeloid leukemia sources. Stem Cells 2007;*

**45** *Takeda A, Goolsby C, Yaseen NR. NUP98-HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34+ hematopoietic cells. Cancer Res 2006;*

**46** *Tonks A, Pearn L, Musson M, et al. Transcriptional dysregulation mediated by RUNX1-RUNX1T1 in normal human progenitor cells and in acute myeloid leukaemia. Leukemia 2007;*

**47** *Verhaak RGW, Goudswaard CS, van Putten W, et al. Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. Blood 2005;*

## Supplementary methods

**Long-term cultures on stroma**. Cord blood (CB) CD34$^+$ cells were derived from neonatal cord blood from healthy full-term pregnancies after informed consent from the Obstetrics departments of the Martini Hospital and University Medical Center in Groningen, The Netherlands. AML blasts from peripheral blood cells or bone marrow cells from untreated patients with AML were studied after informed consent and the protocol was approved by the medical ethical committee. CB and AML mononuclear cells were isolated by density gradient centrifugation and CD34$^+$ cells were selected by MiniMACS. Lentiviral particles were produced as described before.[20,21] Transduction and expansion on MS5 stromal cells in Long Term Culture (LTC) medium are described in the supplemental methods.[20,21] LTC medium (αMEM supplemented with heat-inactivated 12.5% FCS, heat-inactivated 12.5% Horse serum (*Sigma, Zwijndrecht, The Netherlands*), penicillin and streptomycin, 200 mM Glutamine, 57.2 μM β-mercaptoethanol (*Sigma*) and 1 μM hydrocortisone (*Sigma*).

**Single cell assays.** Transduced cells were sorted and cultured in nunclon miniwell trays in IMDM + 10% FCS supplemented with 20ng/ml c-Kit Ligand and IL-3 at 37°C and 5% CO2. They were microscopically analyzed every day for 4-6 days. During these single cells cultures, ~32% of the cells die due to technical reasons, which were discarded from further analysis.

*In vivo* **transplantations into NSG mice.** Before transplantations, mice were sublethally irradiated (2.4 Gy). Following irradiation mice received 3.5g/l neomycin in their drinking water for 2 weeks. The mice were injected with 1.0-2.0x10$^5$ sorted CD34$^+$ cells into the tail vein immediately after transduction. Human cell engraftment was analyzed in the peripheral blood (PB) by flow cytometry after 6-27 weeks of transplantation. For the transplantation of AML cells, GFP$^+$ and GFP$^-$ transduced cells were sorted. Mice transplanted with cells from AML 30, received 300.000 cells/mouse, 40% GFP$^+$ ; AML 31, received 300.000 cells/mouse, 25% GFP$^+$ ; AML 30, received 160.000 cells/mouse, 19% GFP$^+$. Human cell engraftment was analyzed in the peripheral blood (PB) by flow cytometry after 6-9 weeks of transplantation.

**Lentiviral transductions.** AML blasts were transduced in 3 consecutive rounds of 8 to 12 hours with lentiviral supernatant supplemented with 10% FCS, 20 ng/ml IL-3, G-CSF and TPO, and Polybrene (4μg/ml; *Sigma*) and CB CD34$^+$ cells in 2 consecutive rounds of 8 to 12 hours with lentiviral supernatant supplemented with c-Kit Ligand/Flt-3 ligand/TPO (100ng/ml each) and Polybrene (4μg/ml). HL-60 and OCI-M3 cells were transduced in 1 round of 12 hours with lentiviral supernatant supplemented with 10% FCS and Polybrene (4μg/ml).

**Vectors.** A pGL3 basic vector containing the human CITED2 promoter (-3375/+65), followed by a luciferase, was obtained from *Addgene* (plasmid 21486).25 Since this did not contain exon 2, it was named pCt2-Δ5. By means of gDNA PCR the intron and exon 2 were inserted to create pCt2-wt. Both constructs were digested with EcoRI and XmaJI to remove the 3rd PU.1 binding site, creating pCt2-Δ3 and pCt2- Δ3Δ5. The luciferase gene wasubsequently replaced by a d2GFP. Human CITED2 cDNA was obtained through Addgene (plasmid 21487) 26 and cloned into the multiple cloning site of pRRL-SFFV-IRES-tNGFR. A mouse PU.1 cDNA (kind gift of Prof. C Nerlov, *Weatherall Institute of Molecular Medicine, Oxford, UK*) was also cloned into the multiple cloning site of pRRL-SFFV-IRES-tNGFR, either alone or as fusion to the tamoxifen responsive ERt2. pGIPZ shRNAmir against human SPI1, DNMT3A and DNMT3B were obtained from Open Biosystems (*Thermo Scientific, USA*) and the CMV promoter was replaced by an SFFV

promoter. pLKO.1 shRNA against human CITED2, was also obtained from Open Biosystems (The shRNA cassettes were subcloned from the puro based vector to a GFP based pLKO.1 vector, kindly provided by Dr. J. Larsson (*Lund University, Lund, Sweden*). We used 2 different hairpins to minimize off-target effects (*Suppl. Table 3*). We verified whether both hairpins have similar effects so we could use either of them in the leukemic culture assays.

**Streptavidin pull down assay.** Cell extracts intended for streptavidin pull down were made by dissolving cells in HKMG buffer, containing 10mM HEPES (pH 7.9), 100mM KCl, 5mM $MgCl_2$, 10% glycerol 0.5% NP-40, and freshly added 1mM DTT, 0.2mM PMSF and 1 mM of a CLAP cocktail (chymostatin, leupeptin, antipain, pepstatin A) for 10 min. on ice. Lysates were made by sonicating samples five times for 10s with 30s intervals at 10 micron amplitude. Lysates were then spun down at 17,800 × g for 10 min. at 4°C, and the supernatant was collected and stored at -80°C as whole cell extract. Extracts were cleared of aspecific streptavidin binding proteins by incubating with 25µl equilibrated Dynabeads M-280 Streptavidin (*Invitrogen, Life Technologies)* for 1h at 4°C while rotating. Extracts were then cleared of streptavidin beads using a magnetic rack, after which the supernatant was collected and incubated overnight at 4°C with 0.115 µg of a double stranded biotin-labelled oligonucleotide containing either one of five pCt2 PU.1 binding sites or the cJun binding site, in addition to 1 µg of p(dIdC). 25µl of equilibrated Dynabeads was then added and allowed to incubate and rotate for 1h at 4°C. Samples were spun down at 1.000×g for 1min. at 4°C, and 20µl supernatant was collected as "flow through". The pelleted beads were then washed in HKMG buffer five times before finally being dissolved in 20µl HKMG buffer and collected as "pull down". Pull down fractions were then analysed using western blot.

**Antibodies for flow cytometry.** Leukemic LMPP/GMP stain: CD38--FITC (*HIT2, BD*), CD45RA-BrilliantViolet421 (*HI100, Biolegend*), CD90-Biotin (*5E10, Biolegend*), Streptavidin APC-CY7 (*Biolegend*), CD34-APC (*581, BD*), CD123-PECy7 (*6H6, Biolegend*), CD110-PE (*BAH-1, BD*). Lineage PEcy5: CD2 (*RPA-2.10, eBioscience*), CD3 (*UCHT1, eBioscience*), CD4 (*RPA-T4, eBioscience*), CD7 (*124-1D1, eBioscience*), CD8 (*RPA-T8, eBioscience*), CD10 (*MEM78, Biolegend*), CD19 (*HIB19, eBioscience*), CD20 (*2H7, eBioscience*), GPA (*HIR2, Biolegend*). Sort: CD34 PEcy7 (*8G12, BD*), CD38 PE (*HB7, BD*)/APC (*HIT2, Biolegend*), CD271 PE (*C40-1457, BD*)/APC (*ME20.4-1.H4, Milteny*) Myelo-erythroid staining: CD14 APC/PE (*HCD14, Biolegend*), CD15 APC (*HI98, BD*)/PERCP/Pacific Blue (*W6D3, Biolegend*), GPA PE/FITC (*JC159, DAKO*), CD71 AF700 (*MEM-75, Nuclilab*).

**Gene expression profiling.** From 6 independent cord blood batches, $CD34^+$ cells were MACS isolated and transduced with control lentivirus, CITED2 overexpressing lentivirus or a shRNA lentivirus against CITED2. After 2 days transduced $CD34^+CD38^-$ were sorted from each transduction group (Group 1: Control-$CD34^+CD38^-$; Group 2: CITED2-$CD34^+CD38^-$; Group 3: Control-$CD34^+CD38^+$; Group 4: CITED2-$CD34^+CD38^+$; Group 5: Control shSCRC$D34^+CD38^-$; Group 6: shCITED2-$CD34^+CD38^-$; Group 7: Control shSCRC$D34^+CD38^+$; Group 8: shCITED2-$CD34^+CD38^+$. Total RNA was isolated using the RNeasy mini kit from Qiagen (*Venlo, The Netherlands*) according to the manufacturer's recommendations. Q-PCR analysis was used to validate proper overexpression or knock-down of CITED2. 2 CB samples failed this validation and were discarded. RNA from the remaining 4 cord blood was pooled within each group and quality was examined using the Agilent 2100 Bioanalyzer (*Agilent Technologies, Waldbronn, Germany*). Group 3 and group 4 did not pass the quality criteria and were also discarded. On the remaining 6 samples genome-wide expression analysis

was performed on Illumina (*Illumina, Inc., San Diego, CA, USA*) BeadChip Arrays (*Illumina HT12-V4*). Typically, 0.5–1 µg of mRNA was used in labeling reactions and hybridization with the arrays was performed according to the manufacturer's instructions. The expression was quantile normalized using GeneSpring GX software, and from the probesets that were expressed above background (set to 25) subsequent fold differences were calculated. Genes that are positively regulated by CITED2 fulfilled the following criteria: Probesets that are above background indicate a fold change of 2 or more; Upon CITED2 overexpression show an increase in expression; or upon knockdown of CITED2 show a decrease in expression. Genes that are negatively regulated by CITED2 fulfilled the following criteria: Probesets that are above background indicate a fold change of 2 or more; Upon CITED2 overexpression decrease in expression, or upon knockdown of CITED2 increase in expression. The gene expression profiles were ranked based on fold difference and subsequently used in Gene Set Enrichment Analysis. Array data are available at *http://www.ncbi.nlm.nih.gov/geo*, with accession code: *GSE47218*

**DNA isolation, bisulphite conversion and promoter-specific DNA methylation analysis.** Genomic DNA was isolated from 100.000 $CD34^+$ -selected AML cells and bisulphite converted using the EpiTect Plus DNA Bisulfite kit (*Qiagen, Venlo, the Netherlands*) following the manufacturer's recommendations. DNA amplifications were performed on bisulphite-treated DNA using the primers indicated in *suppl. table 3*. Modified DNA wasamplified in a total volume of 25µL. PCR products were cloned into the pJet1.2 vector (*Thermo Scientific*), and 12 randomly picked clones from each AML patient sample was sequenced. Alternatively, the methylation status was examined by pyrosequencing using a Qiagen pyromark Q24 system with 20µl of the PCR product. Pyrosequencing reactions were set up using the Pyromark Gold (*Q24*) Reagent kit (*Qiagen*) according to the manufacturer's instructions. The methylation levels at the target CpGs were evaluated by converting the resulting pyrograms to numerical values for peak heights and expressed as the mean percentage of methylation of all CpGs analysed in the CITED2 promoter sequence.

# Supplementary Tables

## *Supplemental table 1: GSEA results*

| | Gene sets induced by CITED2 in CD34+CD38- HSCs | NES | FDR |
|---|---|---|---|
| 1 | ELVIDGE_HYPOXIA_BY_DMOG_UP | 0,45 | 0 |
| 2 | ELVIDGE_HIF1A_AND_HIF2A_TARGETS_DN | 0,5 | 0 |
| 3 | ELVIDGE_HYPOXIA_UP | 0,41 | 0 |
| 4 | ELVIDGE_HIF1A_TARGETS_DN | 0,51 | 0 |
| 5 | MENSE_HYPOXIA_UP | 0,41 | 0 |
| 6 | KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION | 0,37 | 0 |
| 7 | ROZANOV_MMP14_TARGETS_UP | 0,3 | 0 |
| 8 | GRAHAM_CML_QUIESCENT_VS_NORMAL_DIVIDING_UP | 0,49 | 0 |
| 9 | ODONNELL_TARGETS_OF_MYC_AND_TFRC_UP | 0,41 | 0 |
| 10 | MANALO_HYPOXIA_UP | 0,32 | 0 |
| 11 | CHEN_LVAD_SUPPORT_OF_FAILING_HEART_UP | 0,37 | 0 |
| 12 | MCLACHLAN_DENTAL_CARIES_DN | 0,28 | 0 |
| 13 | MARSON_FOXP3_TARGETS_UP | 0,41 | 0 |
| 14 | TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_10D_UP | 0,29 | 0 |
| 15 | MCLACHLAN_DENTAL_CARIES_UP | 0,27 | 0 |
| 16 | SEMENZA_HIF1_TARGETS | 0,52 | 0 |
| 17 | PAPASPYRIDONOS_UNSTABLE_ATEROSCLEROTIC_PLAQUE_UP | 0,42 | 0 |
| 18 | MISSIAGLIA_REGULATED_BY_METHYLATION_UP | 0,33 | 0 |
| 19 | HOOI_ST7_TARGETS_DN | 0,38 | 0,001 |
| 20 | NAGASHIMA_NRG1_SIGNALING_UP | 0,27 | 0 |
| 21 | KINSEY_TARGETS_OF_EWSR1_FLI1_FUSION_DN | 0,24 | 0 |
| 22 | CYTOSOLIC_PART | 0,56 | 0 |
| 23 | ODONNELL_TFRC_TARGETS_UP | 0,23 | 0 |
| 24 | LINDSTEDT_DENDRITIC_CELL_MATURATION_B | 0,47 | 0,001 |
| 25 | LEONARD_HYPOXIA | 0,48 | 0,001 |
| 26 | VERHAAK_AML_WITH_NPM1_MUTATED_UP | 0,28 | 0 |
| 27 | WINTER_HYPOXIA_METAGENE | 0,25 | 0 |
| 28 | TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_HSC_DN | 0,24 | 0 |
| 29 | HARRIS_HYPOXIA | 0,36 | 0 |
| 30 | AMIT_SERUM_RESPONSE_40_MCF10A | 0,48 | 0,001 |
| 31 | RUIZ_TNC_TARGETS_UP | 0,28 | 0 |
| 32 | ONDER_CDH1_TARGETS_1_UP | 0,28 | 0,002 |
| 33 | REACTOME_PEPTIDE_LIGAND_BINDING_RECEPTORS | 0,4 | 0,003 |
| 34 | BILD_HRAS_ONCOGENIC_SIGNATURE | 0,24 | 0 |
| 35 | KEGG_METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450 | 0,5 | 0,002 |
| 36 | ICHIBA_GRAFT_VERSUS_HOST_DISEASE_35D_UP | 0,29 | 0 |
| 37 | GRAHAM_NORMAL_QUIESCENT_VS_NORMAL_DIVIDING_UP | 0,37 | 0,003 |
| 38 | LIU_VAV3_PROSTATE_CARCINOGENESIS_UP | 0,38 | 0,003 |
| 39 | AMIT_EGF_RESPONSE_60_MCF10A | 0,45 | 0,002 |
| 40 | BROWNE_INTERFERON_RESPONSIVE_GENES | 0,35 | 0,003 |
| 41 | INTERCELLULAR_JUNCTION | 0,55 | 0,001 |
| 42 | GRAHAM_CML_DIVIDING_VS_NORMAL_QUIESCENT_DN | 0,3 | 0 |
| 43 | GRAHAM_CML_QUIESCENT_VS_NORMAL_QUIESCENT_DN | 0,38 | 0,003 |
| 44 | MARTORIATI_MDM4_TARGETS_FETAL_LIVER_UP | 0,28 | 0,001 |
| 45 | REACTOME_GPCR_LIGAND_BINDING | 0,3 | 0,001 |
| 46 | GOLDRATH_IMMUNE_MEMORY | 0,38 | 0,002 |
| 47 | FLECHNER_BIOPSY_KIDNEY_TRANSPLANT_REJECTED_VS_OK_UP | 0,31 | 0,001 |
| 48 | VECCHI_GASTRIC_CANCER_EARLY_DN | 0,26 | 0 |
| 49 | WEINMANN_ADAPTATION_TO_HYPOXIA_DN | 0,51 | 0,003 |
| 50 | KOBAYASHI_EGFR_SIGNALING_24HR_UP | 0,33 | 0,003 |

| | Gene sets repressed by CITED2 in CD34+CD38- HSCs | NES | FDR |
|---|---|---|---|
| 1 | GINESTIER_BREAST_CANCER_ZNF217_AMPLIFIED_DN | -0,31 | 0 |
| 2 | ELVIDGE_HIF1A_TARGETS_UP | -0,42 | 0 |
| 3 | ELVIDGE_HYPOXIA_DN | -0,29 | 0 |
| 4 | ELVIDGE_HIF1A_AND_HIF2A_TARGETS_UP | -0,45 | 0 |
| 5 | WELCSH_BRCA1_TARGETS_1_DN | -0,28 | 0 |
| 6 | ELVIDGE_HYPOXIA_BY_DMOG_DN | -0,4 | 0 |
| 7 | MANALO_HYPOXIA_DN | -0,2 | 0 |
| 8 | REACTOME_CYTOSOLIC_TRNA_AMINOACYLATION | -0,52 | 0 |
| 9 | LIN_SILENCED_BY_TUMOR_MICROENVIRONMENT | -0,42 | 0,001 |
| 10 | OUILLETTE_CLL_13Q14_DELETION_DN | -0,4 | 0,001 |
| 11 | RNA_HELICASE_ACTIVITY | -0,49 | 0,001 |
| 12 | FLOTHO_PEDIATRIC_ALL_THERAPY_RESPONSE_DN | -0,52 | 0,003 |
| 13 | EXOPEPTIDASE_ACTIVITY | -0,52 | 0,001 |
| 14 | DEBIASI_APOPTOSIS_BY_REOVIRUS_INFECTION_DN | -0,21 | 0 |
| 15 | GINESTIER_BREAST_CANCER_20Q13_AMPLIFICATION_DN | -0,22 | 0,001 |
| 16 | FAELT_B_CLL_WITH_VH3_21_UP | -0,35 | 0,002 |
| 17 | NUCLEOLAR_PART | -0,52 | 0,002 |
| 18 | KEGG_PHOSPHATIDYLINOSITOL_SIGNALING_SYSTEM | -0,3 | 0,004 |
| 19 | SIG_CHEMOTAXIS | -0,38 | 0,004 |
| 20 | SOTIRIOU_BREAST_CANCER_GRADE_1_VS_3_UP | -0,21 | 0 |
| 21 | REACTOME_TRNA_AMINOACYLATION | -0,34 | 0,002 |
| 22 | VALK_AML_CLUSTER_4 | -0,42 | 0,008 |
| 23 | KEGG_AMINOACYL_TRNA_BIOSYNTHESIS | -0,33 | 0,004 |
| 24 | CYTOSKELETON_ORGANIZATION_AND_BIOGENESIS | -0,21 | 0,002 |
| 25 | GARY_CD5_TARGETS_DN | -0,15 | 0,001 |
| 26 | NUCLEOLUS | -0,23 | 0,002 |
| 27 | ACTIVE_TRANSMEMBRANE_TRANSPORTER_ACTIVITY | -0,31 | 0,007 |
| 28 | MONNIER_POSTRADIATION_TUMOR_ESCAPE_UP | -0,16 | 0 |
| 29 | PROTEIN_DIMERIZATION_ACTIVITY | -0,22 | 0,004 |
| 30 | ACTIN_CYTOSKELETON_ORGANIZATION_AND_BIOGENESIS | -0,27 | 0,006 |
| 31 | PENG_GLUTAMINE_DEPRIVATION_DN | -0,24 | 0,005 |
| 32 | ACTIN_FILAMENT_BASED_PROCESS | -0,25 | 0,007 |
| 33 | NADERI_BREAST_CANCER_PROGNOSIS_UP | -0,37 | 0,008 |
| 34 | WINNEPENNINCKX_MELANOMA_METASTASIS_UP | -0,19 | 0,005 |
| 35 | YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_11 | -0,23 | 0,008 |
| 36 | FERREIRA_EWINGS_SARCOMA_UNSTABLE_VS_STABLE_UP | -0,2 | 0,002 |
| 37 | KRASNOSELSKAYA_ILF3_TARGETS_DN | -0,39 | 0,008 |
| 38 | NUCLEOBASENUCLEOSIDENUCLEOTIDE_AND_NUCLEIC_ACID_PORT | -0,38 | 0,008 |
| 39 | ATP_DEPENDENT_HELICASE_ACTIVITY | -0,38 | 0,008 |
| 40 | ATPASE_ACTIVITY | -0,22 | 0,006 |
| 41 | SCHLOSSER_MYC_TARGETS_AND_SERUM_RESPONSE_DN | -0,29 | 0,014 |
| 42 | MITSIADES_RESPONSE_TO_APLIDIN_DN | -0,16 | 0,002 |
| 43 | RNA_DEPENDENT_ATPASE_ACTIVITY | -0,45 | 0,009 |
| 44 | BROWNE_HCMV_INFECTION_20HR_UP | -0,18 | 0,002 |
| 45 | TRNA_METABOLIC_PROCESS | -0,41 | 0,008 |
| 46 | SCHLOSSER_MYC_TARGETS_REPRESSED_BY_SERUP | -0,18 | 0,004 |
| 47 | BIOCARTA_RAC1_PATHWAY | -0,43 | 0,011 |
| 48 | ST_INTEGRIN_SIGNALING_PATHWAY | -0,27 | 0,013 |
| 49 | PUJANA_BRCA2_PCC_NETWORK | -0,14 | 0,001 |
| 50 | MCBRYAN_PUBERTAL_BREAST_6_7WK_UP | -0,2 | 0,009 |

**Leukemia related signatures**

**Hypoxia related signatures**

***Supplemental table 2: Primers and probe sequences*** *(next page)*

*The PU.1 core consensus binding site is underlined, h=human, m=mouse*

*Supplemental table 2: Primers and probe sequences*

| **Biotinylated Probes** | **Forward sequence 5'- 3'** | **Reverse sequence 5'- 3'** |
|---|---|---|
| prCT2, PU.1 site 1 | ACCTTTCCTTAGTGGTTGCAA**GAGGAA**TTTAGATTAGACATTCAAAAGCT | AGCTTTTGAATGTCTAATCTAAA**TTCCTC**TTGCAACCACTAAGGAAAGGT |
| prCT2, PU.1 site2 | CGGTTTGCCGGAGAGCCTGGAA**GAGGAA**AGAAGACAAGACACAAAGTAAC | GTTACTTTGTGTCTTGTCTTCT**TTCCTC**TTCCAGGCTCTCCGGCAAACCG |
| prCT2, PU.1 site 3 | GCTCCTGTCCTTGAAAGGAGTG**GAGGAA**AAAAATGCATCTACAAGCGGTG | CACCGCTTGTAGATGCATTTTT**TTCCTC**CACTCCTTTCAAGGACAGGAGC |
| prCT2, PU.1 site 4 | TCCGGGTCTTTAAAACCCAACTC**GAGGAA**GCACAGCCATTCTTCGCTGCC | GGCAGCGAAGAATGGCTGTGC**TTCCTC**GAGTTGGGTTTTAAAGACCCGGA |
| prCT2, PU.1 site 5 | CCTCCGAGGGCAGGGTTTGCCCG**GAGGAA**GAGCGTGAGGCGAAGCTGGGG | CCCCAGCTTCGCCTCACGCTC**TTCCTC**CGGGCAAACCCTGCCCTCGGAGG |
| prJUNB | GCCGCATCCTGCGCCTCGGAA**GAGGAA**CCTGGGCGTGCGAGCCCCCGCGG | CCGCGGGGGCTCGCACGCCCAGG**TTCCTC**TTCCGAGGCGCAGGATGCGGC |

| **Q-PCR primers** | **Forward sequence 5'- 3'** | **Reverse sequence 5'- 3'** |
|---|---|---|
| hCITED2 | CTTTGCACGCCAGGAAGGTC | CTTTGCACGCCAGGAAGGTC |
| hSPI1 | GCGACCATTACTGGGACTTC | ATGGGTACTGGAGGCACATC |
| mSfpi1 | GCCTCAGTCACCAGGTTTCC | GCCTCAGTCACCAGGTTTCC |
| hNCF4 | TCCTCCTCAGTCGGATCAAC | TCCTCCTCAGTCGGATCAAC |
| hCSF3R | AGGCTACCCTCCAGCCATAC | ACGCAGTCCAGGATGGAGTC |
| hB2M | TGGGATCGAGACATGTAAGC | TGGGATCGAGACATGTAAGC |
| hHPRT1 | TTGCTGACCTGCTGGATTAC | TTGCTGACCTGCTGGATTAC |
| hRPL27 | TCCGGACGCAAAGCTGTCATCG | TCCGGACGCAAAGCTGTCATCG |
| hUBC | TAGTCCCTTCTCGGCGATTC | TAGTCCCTTCTCGGCGATTC |
| hGAPDH | CACCCACTCCTCCACCTTTG | CCACCACCCTGTTGCTGTAG |
| hDNMT3A | GGTTCGGAGACGGCAAATTC | CATGGGCTGCTTGTTGTACG |
| hDNMT3B | ACTCAGCCACCTCTGACTAC | CATCTTCCAGGCTGCTCTTG |
| hDDIT4 | CCCTGGACAGCAFCAACAGT | CTCGAAGTCGGGCAACGACA |
| hCDKN1A | ACTAGGCGGTTGAATGAGAG | AGGAAGTAGCTGGCATGAAG |
| hCDKN1B | GCAACCGACGATTCTTCTAC | GTCCATTCCATGAAGTCAGC |
| hCDKN1C | AAGAGATCAGCGCCTGAGAAG | TGGGCTCTAAATTGGCTCACCcc |

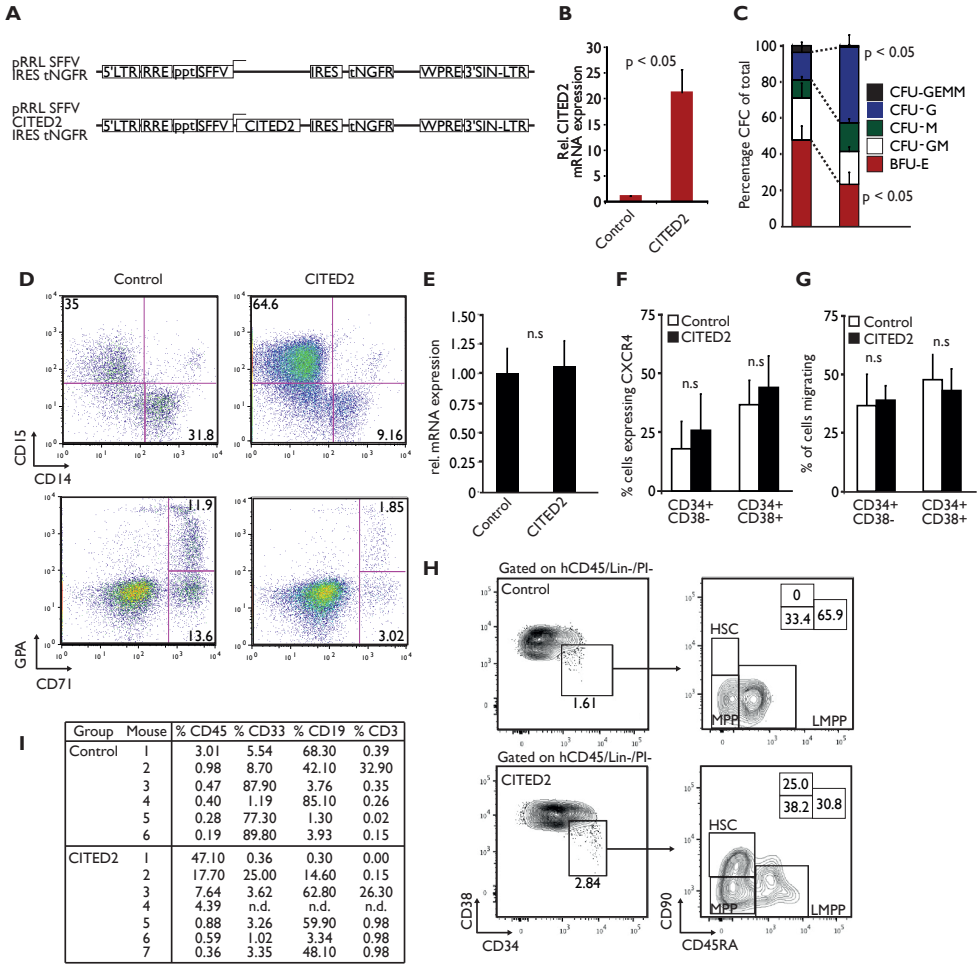
| **ChIP primers** | **Forward sequence 5'- 3'** | **Reverse sequence 5'- 3'** |
|---|---|---|
| prCT2, -10 kb | CCACTATGGCACCAGGTAATC | GGAGGTAGGGACTACTTACAC |
| prCT2, PU.1 site 1 | GTGACCTTTCCAGGGATTTC | GAAGGGGCCACTTCCTATAA |
| prCT2, PU.1 site2 | GTGTCATGTGCCCGTATCT | CCCCTCTGTCTTTTCCTAGTC |
| prCT2, PU.1 site 3 | CCAGCTCCTGTCCTTGAAA | GACCCGGACAAATAATGCG |
| prCT2, PU.1 site 4 | AGGGCGCATTATTTGTCCG | CGGTTCACCTAGCGATCTG |
| prCT2, PU.1 site 5 | GTCTTCGGAGCAGAAATC | AAATCAGCCCTCCTCATC |

| **shRNA-shRNAmir** | **Mature antisense** | **ID number** |
|---|---|---|
| pLKO.1-GFP Control | TTGGTGCTCTTCATCTTGTTG | NA |
| pLKO.1-GFP shCITED2 1 (637) | AAATCCGGCATGTAGTGGTTG | TRC 0000015654 |
| pLKO.1-GFP shCITED2 2 (520) | ATGAACTGGGAGTTGTTAAAC | TRC 0000015655 |
| pGIPZ SFFV Control | ACGTGACACGTTCGGAGAAG | NA |
| pGIPZ SFFV shRNAmir SPI1 | GGTAGGTCATCTTCTTGCG | V3THS_372921 |
| pGIPZ SFFV shRNAmir DNMT3A | ATGCTTCTGTGTGACGCTG | V2LHS_202453 |
| pGIPZ SFFV shRNAmir DNMT3B | GGTAGGTCATCTTCTTGCG | V2LHS_77234 |
| pHR TRIP Control | TTGGTGCTCTTCATCTTGTTG | NA |
| pHR TRIP shCDKN1A | CTCGGTGACAAAGTCGAAG | NA |

| **Sequence primers** | **Forward sequence 5'- 3'** | **Reverse sequence 5'- 3'** |
|---|---|---|
| DNMT3A | GTCCGCAGCGTCACACAGAA | GTGTCGCTACCTCAGTTTGC |

| **Pyrosequence primers** | **Forward sequence 5'- 3'** | **Reverse sequence 5'- 3'** |
|---|---|---|
| PCR amplification | GGAGTGGAGGAAAAAAATGTATTTA | Bio-ACAAAAAAAAAATCCTCCCCTAC |
| Pyrosequencing | ATGAAAATATAAAGGTATAGTTTAG | |
| | GGTATAAGGGTATTTTGGA | |
| | GTGGGGTAGATTTAGT | |

## Supplementary figures

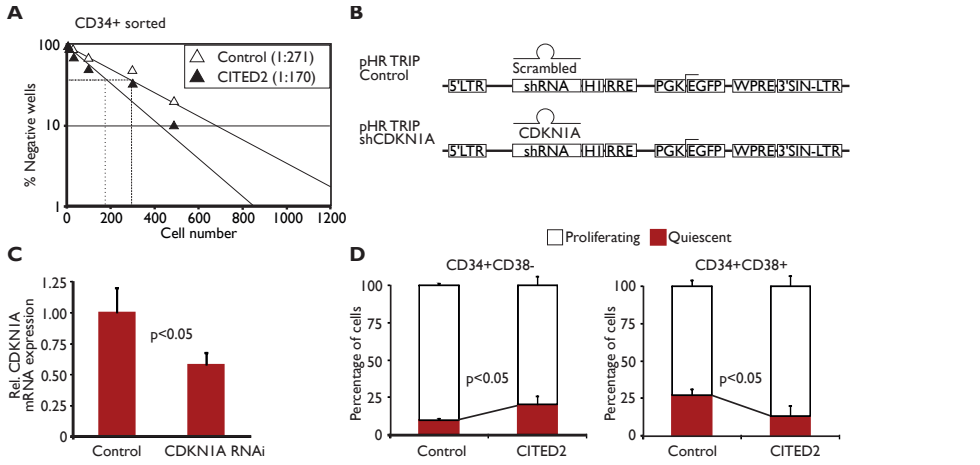



I

| Group | Mouse | % CD45 | % CD33 | % CD19 | % CD3 |
|---|---|---|---|---|---|
| Control | 1 | 3.01 | 5.54 | 68.30 | 0.39 |
| | 2 | 0.98 | 8.70 | 42.10 | 32.90 |
| | 3 | 0.47 | 87.90 | 3.76 | 0.35 |
| | 4 | 0.40 | 1.19 | 85.10 | 0.26 |
| | 5 | 0.28 | 77.30 | 1.30 | 0.02 |
| | 6 | 0.19 | 89.80 | 3.93 | 0.15 |
| CITED2 | 1 | 47.10 | 0.36 | 0.30 | 0.00 |
| | 2 | 17.70 | 25.00 | 14.60 | 0.15 |
| | 3 | 7.64 | 3.62 | 62.80 | 26.30 |
| | 4 | 4.39 | n.d. | n.d. | n.d. |
| | 5 | 0.88 | 3.26 | 59.90 | 0.98 |
| | 6 | 0.59 | 1.02 | 3.34 | 0.98 |
| | 7 | 0.36 | 3.35 | 48.10 | 0.98 |

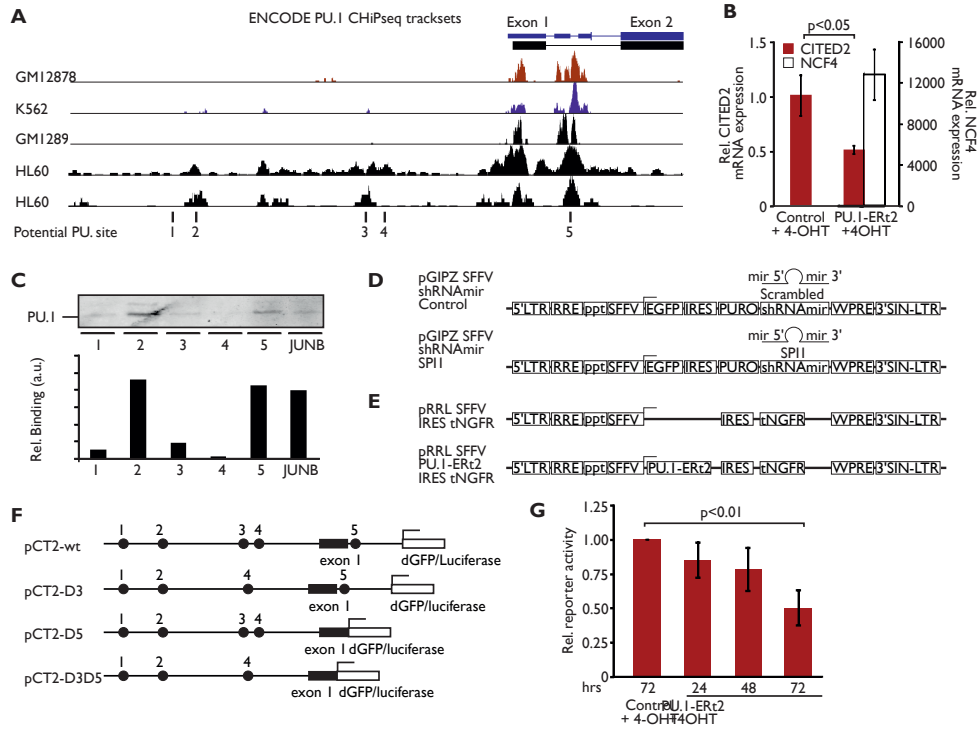
### *Supplemental figure 1*

***A**) Schematic representation of the lentiviral CITED2 overexpression vector. **B**) Representative example of CITED2 overexpression in CD34⁺ CB cells (N=8) **C**) Colony forming cell assay for control or CITED2 transduced CD34⁺ cells. To highlight differences, CFC numbers have been normalized to control and dotted lines have been added to emphasize same colony types. (N=8), error bars denote standard deviation. **D**) Typical FACS plots showing myeloid (upper panels: CD15/CD14, granulocytic vs. monocytic) differentiation at week 4 of culture on MS5 stroma cells. The lower panels show typical FACS plots for erythroid (CD71/GPA double positive cells), differentiation at week 2 of culture on MS5 stroma cells. (N=4) **E**) Relative CXCR4 mRNA expression after CITED2 overexpression (N=4) **F**) CXCR4 protein expression as measured by FACS on CD34⁺CD38⁻ HSCs and CD34⁺CD38⁺ progenitors (N=4), error bars denote standard deviation. **G**) A 24-hrs migration assay of CITED2-transduced HSCs or progenitors towards 100 ng/ml SDF-1 (N=3, error bars denote standard deviation). **H**) FACS plots showing in vivo HSC, MPP and LMPP populations at 28 weeks after transplantation of CITED2 overexpressing cells. Frequencies are depicted in the upper right corner. **I**) Week 24 transplantation data. CD33 myeloid cells, CD19 B-cells and CD3 Tcells are indicated as percentages of hCD45 positive cells.*

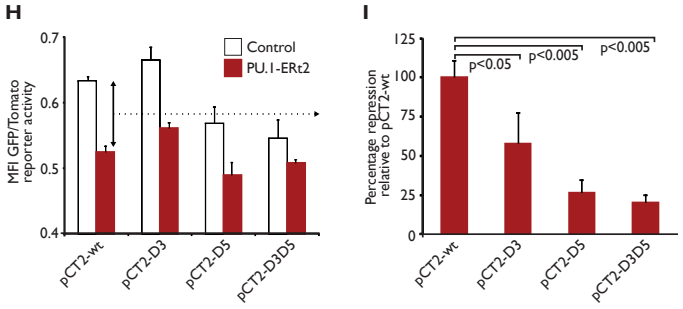



***Supplemental figure 2***

***A**) Limiting dilution analysis of transduced CD34⁺cells. (N=4). The average LTC-IC frequency is indicated in brackets. **B**) Schematic representation of lentiviral H1 driven short hairpin RNA vectors. **C**) Representative example of downregulation of CDKN1A (N=3). **D**) CB CD34⁺ cells were lentivirally infected with control or CITED2. 120 single CD34⁺CD38⁻ HSCs and CD34⁺CD38⁺ progenitors were subsequently sorted into terasaki plates in IMDM plus 20% FCS and 10 ng/ml IL-3 and SCF.After 4 days of culture, each well was microscopically analyzed for cells that had divided or cells that had not divided (N=3). Individual experiments were normalized to compare. Error bars depict standard deviation.*



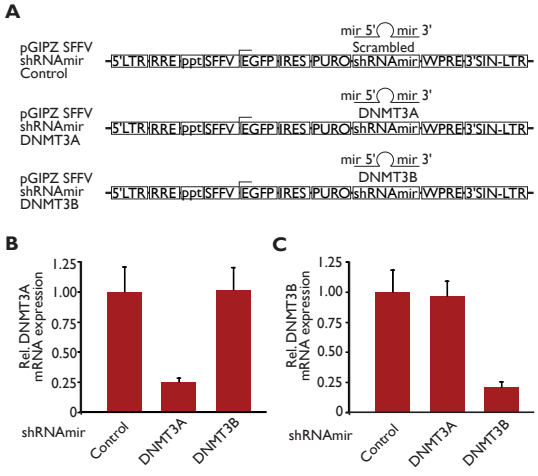

***Supplemental figure 3***




***Supplemental figure 3***

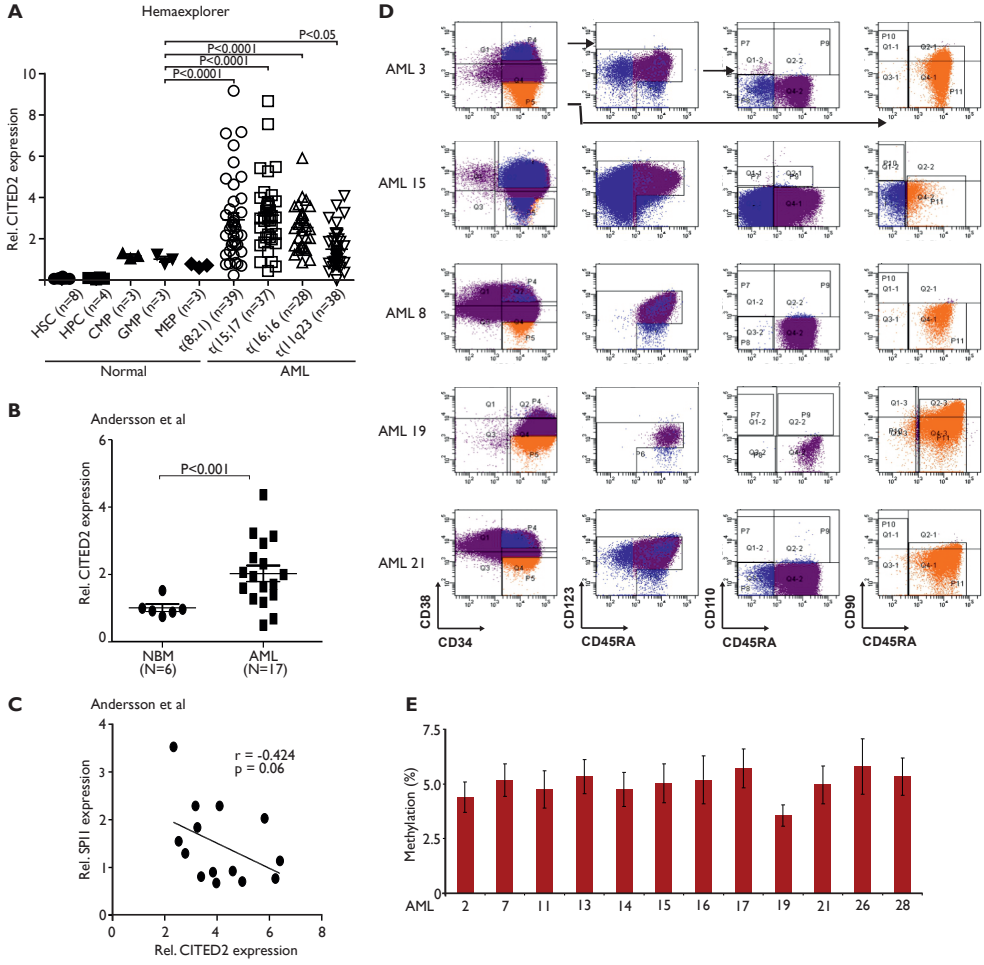
***A**) ChIP-seq track data of PU.1 on the CITED2 promoter in various hematopoietic cell lines obtained from ENCODE. The arrows indicate the potential PU.1 binding sites. **B**) 293T cells were transduced with lentiviral PU.1-ERt2 and stimulated with 4-OHT for 4 days. mRNA was isolated and QPCR for the indicated genes was performed (n=3) **C**) Biotinylated oligos for the putative SPI1 binding sites were synthesized and used for streptavidinmediated pulldown of proteins in cell lysates of 293T cells transduced with PU.1. The upper panel shows a western blot against PU.1, indicating that PU.1 can bind to these putative PU.1 binding sites. A canonical PU.1 binding site from the JUNB promoter was used as a positive control. The lower panel indicates quantification of the western blot, with the binding to the JUNB promoter as the standard set to 1. A typical example is shown (n=2). **D**) Schematic representation of the lentiviral shMir-SPI1 knockdown vector. **E**) Schematic representation of the lentiviral PU.1-ERt2 overexpression vector. **F**) Schematic representation of the CITED2 promoter reporter constructs used. The CITED2 promoter was linked to a GFP or luciferase to create a CITED2 reporter vector (pCt2-wt). Subsequently, PU.1 binding sites were deleted to create pCT2-Δ3, pCT2-Δ5 and pCT2-Δ3Δ5 reporters. **G**) Control and 293T cells stable expressing PU.1 ERt2, were transfected with the pCt2-wt reporter and cells were stimulated with 100 nM 4-OHT. After 4 days either lysates were prepared and luciferase activity was measured (Luciferase vectors) or the Mean Fluorescent Intensity (MFI) of GFP was measured by FACS (GFP vectors). **H**) 293T cells infected with PU.1-ERt2 and control vector were transfected with the pCT2-wt, pCT2-Δ3, pCT2-Δ5 and pCT2-Δ3Δ5 reporters and stimulated with 100 nM 4-OH tamoxifen up to 96 hrs. **I**) Reporter data is normalized to controls and repression of the pCT2-wt reporter (indicated by the arrow in H) was set to 100%.*





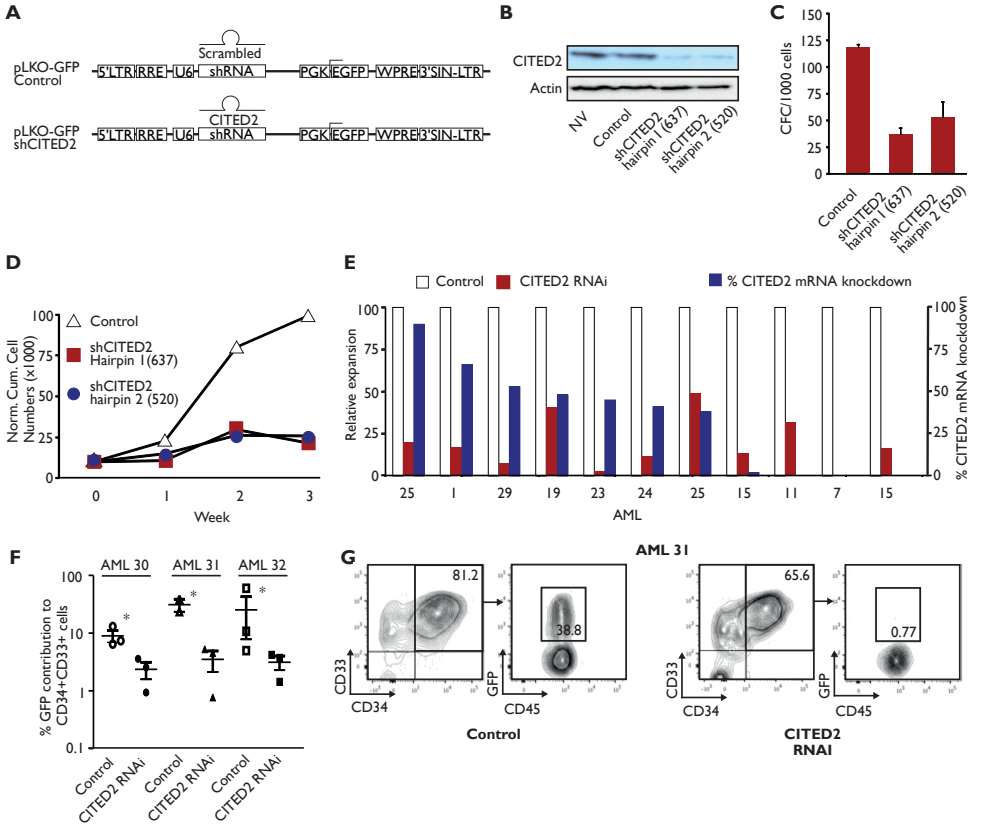

***Supplemental figure 4***

***A**) Schematic representation of the lentiviral shMir-DNMT3A and DNMT3B knockdown vectors. **B**) mRNA expression analysis of DNMT3A knockdown. **C**) mRNA expression analysis of DNMT3B knockdown.*




***Supplemental figure 5***

***A****) Gene expression analysis using the HemaExplorer website. Significant differences in CITED2 gene expression between the various AML groups and the closest related normal counterpart, GMPs, is indicated.* ***B****) Gene expression analysis of CITED2 between CD34⁺ cells from 6 normal bone marrow donors (NBM) and 17 pediatric AMLs.* ***C****) Comparison of CITED2 and SPI1 gene expression in the gene expression arrays from Andersson et al. demonstrating a similar trend in negative correlation between SPI1 and CITED2.* ***D****) FACS analysis of AML samples for CD45RA expression within the CD34⁺CD38⁻ and CD34⁺CD38⁺ populations. For precise antibody staining, see material and methods. (N=5).* ***E****) The methylation status of the CITED2 promoter was examined in CD34⁺ cells from 12 AML patients by pyrosequencing. The methylation levels are expressed as the mean percentage of methylation of all CpGs analysed at the CITED2 promoter sequence with error bars depicting standard deviation.!*




***Supplemental figure 6***

***A**) Schematic representation of the lentiviral U6-driven CITED2 knockdown vectors. **B**) Western blot analysis demonstrating proper knockdown of CITED2 using 2 different hairpins in CB CD34⁺ cells. The numbers indicate the starting target nucleotide. **C**) CFC assay demonstrating that both hairpins against CITED2 have comparable effects on the colony forming potential of CB CD34⁺ cells. **D**) MS5 co-cultures with transduced CB CD34⁺ cells demonstrate that both hairpins show comparable effects on growth kinetics. **E**) Comparison of expansion of CITED2 RNAi cultures relative to its control culture, with the level of CITED2 mRNA knockdown. **F**) GFP⁺ engraftment in NSG mice transplanted with control (n=3/AML) or shCITED2 (n=3/AML) transduced AML CD34⁺ cells. Cells were first gated through human CD45, CD33 and CD34. * indicates p<0.05. **G**) Top: typical FACS plots showing CD33/CD34 double positive cells that contribute to human GFP⁺ engraftment at 9 weeks post transplant in a control-engrafted mouse. Bottom: similar plots demonstrating the absence of GFP⁺ contribution to CD33/CD34 double positive cells after knockdown of CITED2. Cells were first gated through human CD45.*